



Db 588 ---LDTLKLA| | : | | | : : : : :  
588 ---LDTLKLA| | : | | | : : : : :  
588 ---LDTLKLA| | : | | | : : : : :

## RESULT 14

starch synthase (EC 2.4.1.21) precursor - rice  
N/Alternate names: starch synthase  
C:Species: Oryza sativa (rice)  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 16-Aug-2002  
C:Accession: JQ2322; PQ0811  
R:Baba, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shimada, H.; Kobayashi, E.; Ohnishi  
Plant Physiol. 103, 565-573, 1993  
A:Title: Identification, cDNA cloning, and gene expression of soluble starch synthase in  
A:Reference number: JQ2322; MUID:94302151; PMID:7518089  
A:Accession: JQ2322  
A:Molecule type: mRNA  
A:Residues: 1-626 <BAB1>  
A:Cross-references: DDBJ:DL6202; NID:g450484; PIDN:BAA03739.1; PID:g450485  
A:Accession: PQ0811  
A:Molecule type: protein  
A:Residues: 114-129 <BAB2>  
A:Experimental source: seed  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing  
C:Superfamily: starch synthase  
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase  
F:113/Domain: transit peptide (amyloplast) #status predicted <TRS>  
F:114-626/Product: ADPglucose-starch glycosyltransferase, 57kD form #status experimental  
F:122-626/Product: ADPglucose-starch glycosyltransferase, 55kD form #status experimental  
F:440-448/Region: substrate binding #status predicted

Query Match 25.9%; Score 873.5; DB 2; Length 626;  
Best Local Similarity 41.5%; Pred. No. 1.1e-44;  
Matches 197; Conservative 77; Mismatches 154; Indels 47; Gaps 13;

QY 4 IVVAAEAPWSTGGLGVDVTGGLPTELVYKGRHVMVITAPRY-----DOYADADWTSVV 57  
DB 135 VVFVTGEAPYAKSGGLGDVCGSLPTALALGRHVMVMPRYMNGALNNKFNFAFTEKH 194  
QY 58 VDIM---GE-KVRYFHSIKKGVHRVWIDHPWFLAKVWGTGSKLYGRSGADYLDNHRKF 113  
DB 195 IKIPCFGEHEVTFEFHEYRDSVDVWFVDPSY-----HRPGNLVGDNFGA-FGDNQFRY 247  
QY 114 ALFCKAAIEAARVILPGP---GEDCVFVANDHSAIVPVLLKDEYQPGQFTKAKSVLAI 170  
DB 248 TLICYAACAPLILELGGYIYGOKCMFVYNDHSAIVPVLLAAKYPYGYRDAARSLVI 307  
QY 171 HNLAFQGRWEEAFKDTKLPQAAFDKLFATSDGYAKVYVTEATPMEDEKPELTGKTYKKIN 230  
DB 308 HNLAHQGVPEASTYPDLGLPPEWYGALW-----VFPWARHLDKG-----EAVN 354  
QY 231 WLKGGIIADKLVTVPNTATEITAADAAGVELDVTVI--RAKGIEGIVRGMDIEWNPKT 288  
DB 355 FLKGVAVTDRIVTVSQGSWEVTT-AEGGQGLNELLSSRKSVLNGVINGIDINDNWPST 413  
QY 289 DKLSAPYDONSYYAKAAKEALQELGLPVDPTAPLAFICRLEEQKQVDIILAALPK 348  
DB 414 DKFLPHYVSVDLL-SGAKCKAKELQELGLPIRPDVLPIGFTGRLDYQKGDILKLAIPD 472  
QY 349 ILATPKVQIAILGTGKAAVEKLVNAICTYKGRAGKWFPSAPLAHMLTAGADFMVPSR 408  
DB 473 LM-RDNIQFVMLSGDGFEGECWNRSTESGYRDKFRGWGFVSPVSHRTAGCDLILMPSR 531  
QY 409 FEPCCGLIQLHAMHYGTVPVVASTVGGILVDTVK-----EGVTGFMHGALNPDK 454  
DB 532 FEPCCGLNQLYAMGYGVPPVHHGGGLRDVTFENFPAEKGEQGTGWAFSPPLTTEK 586

## RESULT 15

starch synthase (EC 2.4.1.21) isoform STSII-1 - maize (fragment)  
 N:Alternate names: starch synthase isoform STSII-1  
 T01208  
 RESOUR 15

C:Species: Zea mays (maize)  
C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Aug-2002  
C:Accession: T01208  
R:Knight, M.E.; Hahn, C.; Lilley, C.E.R.; Guan, H.P.; Singletary, G.W.; Mu-Forester, Plant J. 14, 613-622, 1998  
A:Title: Molecular cloning of starch synthase I from maize (W64A) endosperm and expro  
A:Reference number: Z14279; MUID:9834055; PMID:9675904  
A:Accession: T01208  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-732 <N1>  
A:Cross-references: EMBL:AF019296; NID:g2811133; PIDN:AAD13341.1; PID:g2655029  
A:Experimental source: strain W64A; endosperm  
C:Genetics:  
A:Gene: SSIIa  
C:Function:  
A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produc  
C:Superfamily: starch synthase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.9%; Score 872; DB 2; Length 732;  
Best Local Similarity 39.7%; Pred. No. 1.7e-44;  
Matches 211; Conservative 77; Mismatches 178; Indels 66; Gaps 15;

QY 2 LDIVMTAAEVAWPSKGGGLDVTGGLPIELVNRGHRVMTIAPRQYADADTSV--VVD 59  
Db 241 MNVIVAAECSPWCKTGGGLDGVVGVGALPKALARGRHVVVPRYGVYEAFDMGIRKYYK 300  
QY 60 IMGE--KVRYFHSIKGVHVRWIDHPWFLAK---VWKGTSKLYGPRSGADYLDNHHKFA 114  
Db 301 AAGODLEVNVFHAFIDGVDFVFIDAPLFHRQDDIYGGSKQIM-----KRMI 348  
QY 115 LFKCAATEAARVLP-----FGCEDCVFVANDWHSALVPVLLKDEYQPKQFTAKSVLA 169  
Db 349 LFKCAVEVPHVPCGVCYGDG-NLVFTANDWHTALLPVYLKAYYRDHGLMQYTFESVLV 407  
QY 170 IHNIATQGRMWEAFKDTKLPPQAADFKLAFSPGAYAKVYTEATPMDEEKPLTGKTYKKI 229  
Db 408 IHNIHQGRPVDPEFYMOLPEHYLQHFLYD-----PVGG---EHA 446  
QY 230 NNLKGGIIADKLVTYSPNKVATIAADAAGVELDVIIRAKG--LEGIVNGMDIEWNP 287  
Db 447 NIFAAGLKMADRVTVTSRGLVWELKT-VGGGGLDIIIRSDNWKINGIVNGIDHQRWNP 505  
QY 288 TDFLSAP----YDQNSVYAGKAAAEALQAEGLGPEVDPTAPLFAFIGRLEEQGVDIIL 343  
Db 506 VDVELRSDGYNTSYLETLDAGKQCKAALQRELGLEVRDVPVLLGFIGRLDGQKGVDIIG 565  
QY 344 AALPKILATPKVOJAILGTGKAAYEKLVNAIGTKYKGRAGVVKVKSAPLAHMLTAGADPM 403  
Db 566 DAMPWI-AGODVQLVMGLGTGRDLERMLQHLREHENKVRGVWGVSPVMAHRTAGADVL 624  
QY 404 LVP-SRPEPCGLIOLHAMHYGTVPVVAOSTGGLDVTKE----GVTGFHMGALNPKDLDEAD 459  
Db 625 VMS-SRPEPCGLNOLYAMVGTVPVVAHVGGLRDTVAPDFPGDAGLGW-----TFDRAE 678  
QY 460 ADALAAATVRRASVFA--GGRYPEMVANCIQDLSWSKPAQKWEGGLEEVVY 509  
Db 679 ANKLIDARHCLDTRKYGESWSKLSOARGMSCDLSMDHAAELYEDVILVAKY 730

Search completed: June 4, 2003, 14:57:26  
Job time : 20.8268 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:50:41 ; Search time 8.74119 Seconds  
(without alignments)  
3088.950 Million cell updates/sec

Title: US-09-980-771A-3\_COPY\_58\_708  
Perfect score: 3370  
Sequence: 1 ALDIVVAAAEVAPWSKTGL.....SASKTSAAKPLVSRATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481.5	44.0	608	1	UGST_MANES
2	1457.5	43.2	607	1	UGST_SOLTU
3	1455.5	43.0	608	1	UGST_ANTWA
4	1415.5	42.0	603	1	UGST_PEA
5	1412.5	41.9	608	1	UGST_IPOBA
6	1397.5	41.5	609	1	UGST_ORYSA
7	1395.5	41.4	603	1	UGST_HORVU
8	1394.5	41.4	609	1	UGST_ORYGL
9	1362	40.4	615	1	UGST_WHEAT
10	1355.5	40.2	608	1	UGST_SORBI
11	1353.5	40.2	605	1	UGST_MATZE
12	902.5	26.8	641	1	UGS2_SOLTU
13	873.5	25.9	626	1	UGS2_ORYSA
14	871	25.8	752	1	UGS3_PEA
15	869	25.8	610	1	UGS3_WHEAT
16	847.5	25.1	788	1	UGS3_SOLTU
17	684.5	20.3	477	1	GLGA_STRPN
18	677.5	20.1	484	1	GLGA_BACSU
19	670.5	19.9	485	1	GLGA_BAGST
20	669	19.9	478	1	GLGA_LACLA
21	653.5	19.4	480	1	GLG1_RHIME
22	652	19.3	480	1	GLGA_AGR5
23	651	19.3	486	1	GLGA_THEME
24	646.5	19.2	480	1	GLGA_RHTR
25	637	18.9	477	1	GLGA_CLOAB
26	622	18.5	482	1	GLGA_CLOPE
27	622	18.5	484	1	GLGA_VIBCH
28	614.5	18.2	476	1	GLGA_BACHD
29	604.5	17.9	486	1	GLG2_RHIME
30	604	17.9	476	1	GLGA_YERPE
31	602.5	17.9	481	1	GLGA_RHILU
32	587	17.4	477	1	GLGA_ECOLI
33	587	17.4	477	1	GLGA_SALTU

## RESULT 1

```

UGST_MANES
ID UGST_MANES STANDARD; PRT; 608 AA.
AC Q43784;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY OR GBSS.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. M.COL 22; TISSUE=Tuberous root;
RX MEDLINE=94083565; PubMed=8260633;
RA Salehuzzaman S.N., Jacobsen E., Visser R.G.F.;
RT "Isolation and characterization of a cDNA encoding granule-bound
RT starch synthase in cassava (Manihot esculenta Crantz) and its
RT antisense expression in potato.";
RL Plant Mol. Biol. 23:947-962(1993).
CC -! FUNCTION: RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE IN RESERVE
CC STARCH.
CC -! CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =
CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
CC -! PATHWAY: Starch biosynthesis
CC -! SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -! TISSUE SPECIFICITY: SYNTHESIZED IN A NUMBER OF DIFFERENT ORGANS,
CC BUT MOST ABUNDANTLY IN TUBERS.
CC -! SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74160; CAA52273.1;
CC Interpro; IPR001296; Glycos_transf_1.
CC Pfam; PF00534; Glycos_transf_1; 1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Glycogen biosynthesis; Transferase; Glycogen biosynthesis.
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT 1 78 CHLOROPLAST [BY SIMILARITY].
CC CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 608 AA; 66968 MW; C9C970CD3011BDBB CRC64;

```

Query Match 44.0%; Score 1481.5; DB 1; Length 608;  
Best Local Similarity 55.0%; Pred. No. 2.4e-81;  
Matches 299; Conservative 64; Mismatches 148; Indels 33; Gaps 9;















```

RN RP SEQUENCE FROM N.A.
RC STRAIN-CV. 12311; TISSUE=Seed;
RA Hsing Y.C., Liu C., Yu H., Hsieh J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
CC -|- PATHWAY: Starch biosynthesis.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U23945; AAC49804.1; --
CC InterPro: IPR001296; Glycos_transf_1.
CC Pfam: PF00534; Glycos_transf_1; 1.
CC KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC FT Transit peptide; Chloroplast; Starch biosynthesis.
CC FT CHAIN 78 608 CHLOROPLAST [BY SIMILARITY].
CC FT BINDING 97 97 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC FT BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 608 AA; 66074 MW; C3133FEA87D2D8A6 CRC64;
CC
CC Query Match 40.2%; Score 1355.5; DB 1; Length 608;
CC Best Local Similarity 50.6%; Pred. No. 8.2e-74;
CC Matches 276; Conservative 76; Mismatches 154; Indels 39; Gaps 11;
CC
CC QY 2 LDIVMVAEAVPWSKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTISVVVDI- 60
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 82 MNVVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVSPRYDOYKDAWDTISVSEIK 141
CC QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 61 MG---EKVRYFHSIKGVHRVWIDHPFLAKVWGKTKGLYGRSGADYLDNHRKRALFC 117
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 142 MGDGYETVRFHCYKRGVDRVFDHPLFLERVWGKTEEKIYGPDACTGDKDNLRFSLIC 201
CC QY |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 118 KAAIEAARVL-----PF--GP-GEDEVFVANDHWSALVPLVLLKDEYQPKGQFTKAKSVLA 169
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 202 QAALAPRILSLNNNNYFSPGYEDVFCVNDHWTGTLPSYLSKNTQSGIYKDAKTAFC 261
CC QY 170 IINIAPQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYTEATPMEEDEKPLTGKTKYKI 229
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 262 IHNISYQGRFAFSDFPELNLPERKSSDFIDGYEK-----PVEG---RKI 304
CC QY 230 NWLKGIIAADKLVTVSPNYATEIAADAGGVELDTVIRAKGIEGIVNGMDIEWNPKT 289
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 305 NWMKAGILEADRLVTSPIYAEGLISGARGELDNIMRLTGITGIVNGMDYSEWDPSKD 364
CC QY 290 KFLSAPYDONSYYAGKAAKALQAEGLPVPDPTAPLFAFICRLEOKGVDTIILALPKI 349
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 365 KYIAVKYDVSTAVEAKALNKEALQAEVGLPVDKPLVAFICRLEOKGVDPVMAAIP-L 423
CC QY 350 LATPKVQVAILTGKAAEKLVAIGTKYKRAKGVKFSAPLAHMLTAGADEMLVPSRF 409
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 424 LMEEDIQIVLLTGKTKKFERMLMSAEKYPDKVRAVVFNAALAHIMAGADLLAVTSF 483
CC QY 410 EPCGLIQLHAMHYGVVPVASTGGGLVDTVKEGVTGFHGMALNPD--KLDEADADALAATV 467
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 484 EPCGLIQIQKRYGTFPCASTGGGLVDVTIEGKTGFHGMRLSVDCNVVPEADVKVAITL 543
CC QY 468 FRASEVAGGRYPENMVANCISODLSWSKPAOKWEGLEE--VYVKGK-----VATAKE 520
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 544 KRAIKVCTPAYEENVKNCMIQDLSWKGPKNWENVILLSGVAGGEPGIEEETAPLAKE 603
CC QY 521 EIKVP 525
CC Db :
CC 604 NVAAP 608

```

```

RESULT 11
UGST_MAIZE STANDARD; PRT; 605 AA.
AC PO4713;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_taxid=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Kloesgen R.B., Gierl A., Schwarz-Sommer Z., Saedler H.;
RT "Molecular analysis of the waxy locus of Zea mays.";
RL Mol. Gen. Genet. 203:237-244(1986).
CC -|- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -|- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
CC -|- PATHWAY: Starch biosynthesis.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X03935; CAA27574.1; --
CC EMBL: M24258; AAA33520.1; --
CC PIR: S07314; S07314.
CC MaizeDB; 15806; --
CC InterPro: IPR001296; Glycos_transf_1.
CC Pfam: PF00534; Glycos_transf_1; 1.
CC KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC FT Transit peptide; Chloroplast; Starch biosynthesis.
CC FT CHAIN 73 605 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC FT BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 605 AA; 65966 MW; 137F15207DBFC189 CRC64;

```

```

Query Match 40.2%; Score 1353.5; DB 1; Length 605;
Best Local Similarity 50.4%; Pred. No. 1.1e-73;
Matches 275; Conservative 79; Mismatches 153; Indels 39; Gaps 11;
CC
CC QY 2 LDIVMVAEAVPWSKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTISVVVDI- 60
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 77 MNVVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVSPRYDOYKDAWDTISVSEIK 136
CC QY 61 MG---EKVRYFHSIKGVHRVWIDHPFLAKVWGKTKGLYGRSGADYLDNHRKRALFC 117
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 137 MGDGYETVRFHCYKRGVDRVFDHPLFLERVWGKTEEKIYGPVAGTDYRDNLQRLSLC 196
CC QY 118 KAAIEAARVL-----PF--GP-GEDEVFVANDHWSALVPLVLLKDEYQPKGQFTKAKSVLA 169
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 197 QAALAPRILSLNNNNYFSPGYEDVFCVNDHWTGTLPSYLSKNTQSGIYKDAKTAFC 256
CC QY 170 IINIAPQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYTEATPMEEDEKPLTGKTKYKI 229
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 257 IHNISYQGRFAFSDFPELNLPERKSSDFIDGYEK-----PVEG---RKI 299
CC QY 230 NWLKGIIAADKLVTVSPNYATEIAADAGGVELDTVIRAKGIEGIVNGMDIEWNPKT 289
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Matches	212;	Conservative	90;	Mismatches	161;	Indels	69;	Gaps	19;
QY	3	DIWVAAAEVAPSKTGGGLDVTGGPLIELVKGRHVMYTIAPRY-----DOYADAWDTS	55						
Db	132	NIIFTAEAPYSKTGGGLDVGSLPMAAARSHRVMVSPRYLNGGSPDEKYANAVDLD	191						
QY	56	V--VVDIMG--BEKRYFHSIKKGVRHVWTDHPWFLAKVWGKGTGSKLYGPRSGADYLDNHK	111						
Db	192	VRATVHCFGDAQEVAFYHEYRAGVDMVDFDHSYC-----RFGTP-YGDIYGA-FGDNOF	244						
QY	112	REALFCKAAIEARVLPCGP--GDCVFVANDWSALVPLVLLKDYQKQFTAKSVL	168						
Db	245	RTLLSHAAACEAPVLPGLGFTYGEKCLFANDWHAALVPLLLAAKYRYGVYKQARSIV	304						
QY	169	AHNIAFOGRMMEEAFKDTLPQAADFKLAFSDGYAKVYTEATPMEDEDEKPLITGKYTK	228						
Db	305	AHNIAHQGVPAVTNNLGLPPQWYGAV---EWIETWARAHALD-----TGTE---	351						
QY	229	INWLKGIITADKLVTVSPNYATEIAADRAAGGVLEDTVI--RAKGTEGVNGMDIEWNP	286						
Db	352	VNVLKGAIAVADRILTVSOGYSWEITT-PEGGVLGHELLSSROSUVNGTNGIDVNDWP	410						
QY	287	KTDKFLSAPYDONSVTAGKAAAKEALQAEGLGEPDPTAPLFAFICGLEBOKGVDIILAA	346						
Db	411	STDEHTASHYSINDL-SGKVOCKTDLQKELGPIRDPDCLPTIGTIGRLDYQKGVDIILSAI	469						
QY	347	PKLATPKVQIILGTGKAAYEKLVNAIGTKYGRAGVVKFSAFLAHLMTAGADEMLVP	406						
Db	470	PELMQN-DVOVVMVGSGEKQYEDWMRHTENLFKDFRAWGVFNVPVSHRITACDILLMP	528						
QY	407	SRPEPCGLIOLHAMHYGTVPVASTGGLVDIVK-----EGV---TGFFHMGALNPKLD	456						
Db	529	SRPEPCGLIOLHAMHYGTVPVASTGGLVDIVKDFNFDYQEGIGEGTGWTFSPITSEKL-	587						
QY	457	EADADALAAATVRRASEVFAGGRYPF-----MVANCISDLSWSKPAQKWE	501						
Db	588	--LDTLKLAI-----GTYTEHKSWEGLMRRGMDGRDYSWENAAIQYE	627						
RESULT 13									
ID	UGS2_ORYSA	STANDARD;	PRT;	626	AA.				
AC	Q40739;								
DT	01-NOV-1997 (Rel. 35, Created)								
DT	01-NOV-1997 (Rel. 35, Last sequence update)								
DT	15-JUN-2002 (Rel. 41, Last annotation update)								
DE	Soluble glycogen [starch] synthase, chloroplast precursor								
DE	(EC 2.4.1.11) (SSS).								
OS	Oryza sativa (Rice).								
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;								
OC	Ehrhartoideae; Oryzeae; Oryza.								
OX	NCBI_TaxID=4530;								
RN	[1]								
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.								
RC	STRAIN=cv. Japonica; TISSUE=Seed;								
RX	MEDLINE=94302151; PubMed=7518089;								
RA	Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,								
RA	Kobayashi E., Ohnishi S., Tanaka K.-I., Arai Y.								
RT	*Identification, cDNA cloning, and gene expression of soluble starch								
RT	synthase in rice (Oryza sativa L.) immature seeds.*;								
RL	Plant Physiol. 103:565-573(1993).								
CC	-!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =								
CC	UDP + {(1,4)-alpha-D-glucosyl}(N+1).								
CC	-!- PATHWAY: Starch biosynthesis.								
CC	-!- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.								
CC	-!- TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.								
CC	-!- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE								
CC	PURIFIED: RSS1, RSS2 AND RSS3.								
CC	-!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE								
CC	FAMILY.								
-----									
This SWISS-PROT entry is copyright. It is produced through a collaboration									

This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMLB; D16202; BAA03739.1; -
DR	InterPro; IPR001296; Glycos.transf.1.
DR	Pfam; PF00534; Glycos.transf.1.
DR	Glycosyl biosynthesis; transferase; Glycosyltransferase;
KW	Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
FT	TRANSIT 1 113
FT	CHAIN 114 626
FT	CHAIN 122 626
FT	BINDING 147 147
FT	SEQUENCE 626 AA; 68451 MW; 03E4182507D26658 CRC64;
SQ	

Query Match 25.9%; Score 873.5; DB 1; Length 626;  
Best Local Similarity 41.5%; Pred. No. 5.2e-45;  
Matches 197; Conservative 77; Mismatches 154; Indels 47; Gaps 13;

QY	4	IWMVAEAVPWSKTGGGLDVTGGLPIELVKGHRVMTIAPRY-----DOYADAWDTSVV	57
Db	135	VVFVTEGASPYAKSGGLDVGCSLPIALALRGHRVWVMPRYMGALNKNFANAFTYKH	194
QY	58	VDIM--GE-KVYFHSIKKGHRVWIDHDPFLAKWVGKTSKLYGSPRGADYLDNHKFR	113
Db	195	IKIPCFGGEHEVTFPHRYRDSVDWVFDVHPSY-----HRPGNLYGDNFGA-FGDNQFQY	247
QY	114	ALFCKAAATEAARVLPPFG--GECVFFVANDWHSALVPVLKLDYQPKQSOFTKAKSVLAI	170
Db	248	TLLCYAAACEAPLILELGGYIYQKCMFVNDWHSALVPLVLLAAKTRPGVTRDASVLVI	307
QY	171	HNIAFOGRWEEAEKDYKLPOAADFKLAFSDGIYAKVYTEATPMEDEKPLTKGYKKIN	230
Db	308	HNLAHQGVPEASTYDPLDGLPPEWGALEW-----VPPEWARRHALDKG-----EAVN	354
QY	231	WLKGIIAADRKLIVTSVPNYAFETIAADAAGGVELDTVI--RAKGLIEGVINDIEWNKPT	288
Db	355	FLKAVVTVADRIVTVSQYSWEVIT-ABGGOGNELISSRKSVLNGIYNGIDINDWNPST	413
QY	289	DKFLSAPYDQNSVYAGAAAKEALQELGLPVDPTAPLFAFIGRLEQKGVDIILAAFLPK	348
Db	414	DKFLPYHVSDDL-SGRAKCAEAKQELGLPIRPDVPLIGTIGRLDYQKGLDLKLAIPD	472
QY	349	ILATPKVQOIALTGTAKAAYEKLVAIGTKYKGRAGGVVYKFSAPLAHMLTAGADEMLVPSR	408
Db	473	LM-RDNTOFVMLGSDPGFEGWMKSTESGYEDKFRGVWGVSPVSHRITAGCDILLMPSR	531
QY	409	FEPGLQLQHAMHYGVVPPVASTGGLVDTVK-----EGVTGFHMGALNPDK	454
Db	532	FEPCGMLQMLAMQYGVTVVYHGTGGLRDVTENFPFAEKGEQGGTGWAFSPLTIPK	586

	RESULT 14				
D	GS3_PEA				
S	UGS3_PEA	STANDARD;	PRT;	752 AA.	
G	Q43093;				
T	C T				
I	01-NOV-1997 (rel. 35, Created)				
T	01-NOV-1997 (Rel. 35; Last sequence update)				
T	15-JUN-2002 (Rel. 41; Last annotation update)				
E	Glycogen [starch] synthase, chloroplast precursor (EC 2.4.1.11)				
E	(GBSSI) Granule-bound starch synthase II).				
S	pisum sativum (Garden pea).				
C	Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
C	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
X	eurosid I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.				
X	NCBITaxID=3888;				
N	[1]				
P	SEQUENCE FROM N.A., AND SEQUENCE OF 58-73.				
F	STRAIN=cv. BC1/RR; TISSUE=Embryo;				
X	MEDLINE=93251108; PubMed=1302049;				

Dry I., Smith A., Edwards A., Bhattacharyya B., Dunn P., Martin C.;  
 "Characterization of cDNAs encoding two isoforms of granule-bound  
 starch synthase which show differential expression in developing  
 storage organs of pea and potato.";   
 Plant J. 2:193-202(1992).  
 CC -|- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =  
 CC UDP + [(1,4)-alpha-D-glucosyl](N+1).  
 CC -|- PATHWAY: Starch biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: CHLOROPLAST/AMYLOPLAST, SOLUBLE AND GRANULE-  
 CC BOUND.  
 CC -|- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -|- DEVELOPMENTAL STAGE: MOST HIGHLY EXPRESSED IN EARLY EMBRYOS.  
 CC -|- LEVELS DECLINE IN LATER STAGES OF DEVELOPMENT.  
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL-/PLANT GLYCOGEN SYNTHASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

CC	EMBL; X88790; CAA61269.1; ..
DR	InterPro: IPR001296; Glycos_transf_1.
DR	Pfam: PF00334; Glycos_transf_1; 1.
DR	Glycogen biosynthesis; transferase; Glycosyltransferase;
KW	transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
KW	TRANSIT 1 57 CHLOROPLAST.
FT	CHAIN 58 752 GLYCAGEN [STARCH] SYNTHASE.
FT	BINDING 275 275 UDP-GLUCOSE (BY SIMILARITY).
FT	SEQUENCE 752 AA; 83617 MW; E0496420CD359395 CRC64;
SQ	

Query Match	25.8%;	Score 871;	DB 1;	Length 752;
Best Local Similarity	38.2%;	Pred. No. 9.2e-45;		

Matches	209;	Conservative	82;	Mismatches	160;	Indels	96;	Gaps	16;																																														
QY	2	LDIVMVAARF	PWSKTGG	LDYTG	GLPIELV	KRGHRV	TIAPRY	QOYAD	AWT	TSV--	VVD	59																																											
Db	261	MNIIIVSAE	CAPSWKTGG	LDG	VAGSLG	KALARR	HRVMI	VA	PHYGN	YAF	AHDIGV	KRYK 320																																											
QY	60	IMGE--	KVRYPHS	IKKG	VHR	WIDH	EPW	FLAK	VWG	TGSK	LYGPR	SGADYLDN	NKHFALPC 117																																										
Db	321	VAGDM	EVTFY	TDG	VDIV	FDSP-----	IFRN	LES	NIY---	GGN	PLDIL	RRMVL	FC 371																																										
QY	118	KAAIEAAR	VLP-----	PG	GEDC	VP	AND	WISAL	VP	LVLL	KDYQ	PKG	QFT	TKAKS	YLA	HN 172																																							
Db	372	KAAVEP	PHV	PCG	IGY	GDG--	NL	VE	I	AND	W	H	T	A	L	L	PVY	L	K	A	R	D	H	G	L	M	N	T	R	S	V	L	I	N 430																					
QY	173	IAFG	R	M	W	E	E	A	F	K	D	T	K	L	P	Q	A	A	F	D	K	L	A	F	S	D	G	V	A	K	V	T	E	A	T	P	A	T	E	E	K	P	L	T	G	T	K	Y	K	I	N	L 232			
Db	431	IAHQ	R	G	R	G-----	P	V	E	D	N	T	V	L	S	G	N	Y	L	D	L	E	K	M	D	P	V	G	G	H-----	F	N	I	F 469																					
QY	233	KGGI	A	A	D	K	L	V	T	S	P	N	A	T	E	T	A	A	A	A	G	V	E	L	D	T	V	I	R	A	G--	I	E	G	I	V	G	M	D	I	E	E	N	P	K	T	D 290								
Db	470	AAGL	T	A	D	R	I	V	T	S	H	G	A	W	E	L	K	T--	S	E	G	G	L	H	N	I	N	E	S	D	K	F	R	G	I	V	G	V	D	T	K	M	N	P	Q	F	D 528								
QY	291	FLSAP----	X	D	N	S	V	A	G	K	A	A	K	A	L	A	E	L	G	P	D	T	A	P	L	F	A	F	T	G	R	E	E	O	K	V	D	I	L	A	A	L 346													
Db	529	YLT	S	D	G	Y	T	N	L	K	T	L	Q	T	K	R	C	K	A	L	O	R	E	L	G	L	P	V	R	E	D	V	P	T	I	S	T	F	I	R	L	D	H	O	K	V	D	L	I	A	E	A	I 588		
QY	347	P	K	I	L	A	T	P	K	V	O	J	A	I	L	G	T	K	A	Y	E	K	L	V	N	A	I	G	T	K	Y	K	R	A	G	V	K	F	S	A	P	L	A	M	L	T	A	G	F	D	E	L	V	P 406	
Db	589	P	W	M	--	S	H	D	V	L	M	L	G	T	R	A	D	L	E	O	M	L	K	E	F	A	C	H	C	D	K	I	R	S	V	G	F	S	V	K	M	A	H	R	I	T	A	G	S	D	I	L	L	M	P 647
QY	407	S	R	E	P	E	C	L	I	L	H	A	M	H	T	G	T	P	V	W	A	S	T	G	L	V	T	V	K-----	E	G	V	T	G	F	H	M	C	A	L	N	P	D	K	L	E	A	D 460							
Db	648	S	R	E	P	E	C	L	N	O	L	Y	A	M	S	T	G	T	P	V	P	V	H	G	V	G	L	R	D	I	V	P	N	F	E	D	E	S	G	V	G-----	T	F	D	R	A	E 699								
QY	461	D	A	L	A	A	T	V	R	A	S	E	F	V	A	G	R	Y	E	M	V	A	N	C	I-----	S	D	L	S	W	S	K	P	A	Q	K	W	E	G 502																
Db	700	N	K	L	M	A	A-----	L	W	N	C	L	L	T	K	Y	K	K	S	W	E	G	I	O	E	R	G	M	S	O	D	L	S	W	N	A	O	O	Y	E	E 743														



221	DB	VTADRIVTSQGYSEWVTT-AEGQGGLNELLSSRKSVLNGTVNGDINDWNPITDDKCLPH	277
295	QY	PYDONSYYAGKAAAKEALQAEGLPVDPOTAPLFAFIFGRLEQKGVDDIILAALPKLATPK	354
280	DB	HYSDDDL-SGKAKCAELQKELGHPVRDEDVLGIFGRDLYQKGLDKMAIPELM-RED	337
355	QY	VQIAILGTGAAAYEKLNAIGTKYKRAKGVVYKFSAPLAHMITAGADFMLVPSRPEPGL	414
338	DB	VQFVMLGSGDPIFGWWRSTESSYKDKFGWGVGFSVPYSHRITAGCDILLMPSRPEPGL	397
415	QY	IQLHAMHYGVVVPVASTGGLVDVTYK-----BGVTGFHMGALNPDKLDEADADALAA	465
398	DB	NQLYAMQYGTVPVYVHGTTGGLRDTVETENFPFGAKGEEGTGWAFSPITVDKMLWA-	450
466	QY	TVRRASEVFAGGRIPEMVYANCISQDLSNKSQAQKWEGLLLEVVYKGGVATAKKEIKVP	525
451	DB	-LRTAMSTFRE-----HKPS-WEGLM-----KRGMTKDH	477
526	QY	VAEKIPGDLPAVSYAPNTLKPVSASVEGNGAAAPKVGITAPAMGAWRATTPSGPSPAAAT	585
478	DB	TWDHAPSTSRSSSGPSTWNTFSC-----RRGLGRSKCESPALKITSSSS	522
586	QY	---PKVTITYKPALPAKPKTAGLK-LAGEASTTSTSENGAASNGNGNGASKASKYAAAP	641
523	DB	FRGPEGYPCTLRCPATVESQCACLLWFAG-----SRTYDCCA-----AAAVTASGGRQ	570
642	QY	LVSAAATRK	649
571	DB	LQFWGIRK	578

***This Page Blank (uspro)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:35:26 ; Search time 32.3424 Seconds  
(without alignments)  
4147.394 Million cell updates/sec

Title: US-09-980-771a-3\_COPY\_58\_708

Perfect score: 3370  
Sequence: 1 ALDIYVAAEAVPWSKTGGL.....SASKTSAAPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3366	99.9	708	10	O64925 chlamydomon
2	1463.5	43.4	613	10	Q93YB1 pisum sativ
3	1450.5	43.0	604	10	Q9FR03 perilla fru
4	1424.5	42.3	608	10	Q93VD9 ipomoea bat
5	1423.5	42.2	606	10	Q9XIS6 phaseolus v
6	1423.5	42.2	607	10	Q9ZSQ5 astragalus
7	1404	41.7	608	10	Q8VYU1 oryza sativ
8	1397.5	41.5	609	10	Q94LY7 oryza sativ
9	1395.5	41.4	603	10	Q8SA49 hordeum vul
10	1390	41.2	610	10	Q9MAQ0 arabidopsis
11	1389.5	41.2	609	10	Q8S9C4 oryza sativ
12	1381	41.0	605	10	Q9SXX3 triticum ae
13	1380	40.9	605	10	Q9FUF6 triticum ae
14	1379	40.9	599	10	Q9SQS8 triticum ae
15	1376	40.8	606	10	Q43012 oryza sativ
16	1375	40.8	604	10	Q9S7N5 triticum ae

17	1371	40.7	534	10	Q9W2G8	08w2g8 triticum ae
18	1369	40.6	604	10	Q9SLS6	Q9sls6 triticum tu
19	1367.5	40.6	604	10	Q9SLS7	Q9sls7 triticum tu
20	1367	40.6	605	10	Q9SQS1	Q9sq51 aegilops sp
21	1366	40.5	605	10	Q9SLS8	Q9sls8 triticum tu
22	1362.5	40.4	604	10	Q9SXX4	Q9sxk4 triticum ae
23	1360.5	40.4	604	10	Q9SLS9	Q9sls9 triticum tu
24	1354.5	40.2	605	10	Q9SQS2	Q9sq52 triticum mo
25	1353.5	40.2	574	10	Q9SYU0	Q9syu0 triticum ae
26	1263.5	37.5	565	10	Q9XEN9	Q9xen9 triticum ae
27	898	26.6	792	10	Q9MAC8	Q9mac8 arabidopsis
28	895.5	26.6	313	10	Q9LKD3	Q9lkd3 vaquelinia
29	893.5	26.5	313	10	Q9LKE0	Q9lke0 prunus virg
30	893.5	26.5	313	10	Q9FYU8	Q9fyu8 aruncus dio
31	891.5	26.5	647	10	Q9LEB9	Q9leb9 triticum ae
32	891.5	26.5	647	10	Q9SOH0	Q9sqn0 aegilops ta
33	891.5	26.5	647	10	Q9SQG9	Q9sqg9 triticum ae
34	890.5	26.4	647	10	Q9LEF0	Q9leco triticum ae
35	886.5	26.3	313	10	Q9LKE8	Q9lke8 kageneckia
36	886	26.3	576	10	Q64926	064926 chlamydomon
37	885.5	26.3	313	10	Q9LKD1	Q9lkd1 exochorda r
38	884.5	26.2	643	10	Q9M5A3	Q9m5a3 hordeum vul
39	879.5	26.1	313	10	Q9LKF5	Q9lkf5 amelanchier
40	879.5	26.1	313	10	Q9LKD6	Q9lkd6 vaquelinia
41	878.5	26.1	313	10	Q9LKE7	Q9lke7 kageneckia
42	878.5	26.1	313	10	Q9LKE1	Q9lke1 prinsepia s
43	877.5	26.0	313	10	Q9LKE6	Q9lke6 malus sarge
44	877.5	26.0	313	10	Q9LKD8	Q9lkd8 rosa multif
45	876.5	26.0	313	10	Q9LKE5	Q9lke5 oemleria ce

## ALIGNMENTS

RESULT 1  
O64925 PRELIMINARY; PRT 708 AA.  
AC O64925  
DT 01-AUG-1998 (TREMBLrel. 07, Created  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).  
GN STA2  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=137C;  
RA D'Huist C., Wattedled F., Ral J.-P., Abel G.J., Kossmann J.,  
RA Ball S.G.;  
RT "Cloning of a cDNA encoding for the GBSI in the green alga  
RT Chlamydomonas reinhardtii.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Wattedled F., Ball S.G., D'Huist C.;  
RT "Granule-bound starch synthase I: A major enzyme involved in the  
RT biogenesis of B-crystallites in starch granules.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF026420; AAC17969.3;  
DR EMBL; AF433156; AAL28128.1;  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR InterPro; IPR002114; HPR\_Serp\_site.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
KW Glycosyltransferase; Transferase; Transin peptide.  
FT TRANSIT 1 57 POTENTIAL.  
FT CHAIN 58 708 GRANULE-BOUND STARCH SYNTHASE I.  
SQ SEQUENCE 708 AA; 74623 MW; 7D2A5A07D8606469 CRC64;

Query Match 99.9%; Score 3366; DB 10; Length 708;

Best local Similarity 99.8%; Pred. No. 4.5e-190;  
Matches 650; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALDIWVAEAVPWSKTGGLGDTGGLPIELVKRGHRYMTIAPRYDQYADADMTSVVVDI 60  
DB 58 ALDIWVAEAVPWSKTGGLGDTGGLPIELVKRGHRYMTIAPRYDQYADADMTSVVVDI 117

QY 61 MGEKVRYPHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFCRAA 120  
DB 118 MGEKVRYPHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFCRAA 177

QY 121 IEAARVLPFGEGDCVFAVDWHSALVPVLKDEYQPKGQTKAKSVLAHNIAPQGRMW 180  
DB 178 IEAARVLPFGEGDCVFAVDWHSALVPVLKDEYQPKGQTKAKSVLAHNIAPQGRMW 237

QY 181 EEAERKDLKPOAEFKLAFSDGYAKVYTEATPMEDEKPPITGTYKKINWLGKGIAD 240  
DB 238 EEAERKDLKPOAEFKLAFSDGYAKVYTEATPMEDEKPPITGTYKKINWLGKGIAD 297

QY 241 KLVTSPNYATEIAADAAGGVELDTVIRAKIEGIVNGMDIEENPKTKDPLSAPYDONS 300  
DB 298 KLVTSPNYATEIAADAAGGVELDTVIRAKIEGIVNGMDIEENPKTKDPLSAPYDONS 357

QY 301 VYAKAAKALQALQELGLPVDPTAPLFAFGRLEEKGVDIILAAALPKILATPKVQJAIL 360  
DB 358 VYAKAAKALQALQELGLPVDPTAPLFAFGRLEEKGVDIILAAALPKILATPKVQJAIL 417

QY 361 GTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRFPCCGLIQLHAM 420  
DB 418 GTGKAAEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSRFPCCGLIQLHAM 477

QY 421 HYGTVPVVASTGGVLDVTKGVTGFGHMGALNPDKLDEADADALAAATVRRASEVPAGGRYP 480  
DB 478 HYGTVPVVASTGGVLDVTKGVTGFGHMGALNPDKLDEADADALAAATVRRASEVPAGGRYP 537

QY 481 EMVANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKETPGDLPVASYA 540  
DB 538 EMVANCISQDLSWSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKETPGDLPVASYA 597

QY 541 PNTLKPVASVEGNGAAAPKVGTTAPAMGAWRATTSGSPSAAATPKVTTYKPALPATAK 600  
DB 598 PNTLKPVASVEGNGAAAPKVGTTAPAMGAWRATTSGSPSAAATPKVTTYKPALPATAK 657

QY 601 PTKAGLKLAGEASTTSTSENGASNGNGASASKTSAKPLVSAATRKA 651  
DB 658 PTKAGLKLAGEASTTSTSENGASNGNGASASKTSAKPLVSAATRKA 708

RESULT 2

Q93YB1 PRELIMINARY; PRT; 613 AA.

AC Q93YB1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Granule-bound starch synthase (EC 2.4.1.11).  
GN GBSSI.  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. OKDONG; TISSUE=DEVELOPING SEEDS;  
RA Edwards A., Vincken J.P., Visser R., Zeeman S., Smith A.M., Martin C.;  
RT "Discrete forms of amylose are synthesised by isoforms of GBSSI in pea";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ345045; CAC69955.1;  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR InterPro; IPR001917; NHtransf\_2.  
DR Pfam; PF00534; Glycos\_transf\_1.1.

DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; UNKNOWN\_1.  
KW Glycosyltransferase; Transferrase; Transit peptide.  
FT TRANSIT 1 85  
FT CHAIN 86 613 GRANULE BOUND STARCH SYNTHASE.  
SQ SEQUENCE 613 AA; 67626 MW; 5A653461128F97EB CRC64;

Query Match 43.4%; Score 1463.5; DB 10; Length 613;  
Best Local Similarity 53.0%; Pred. No. 4.1e-173;  
Matches 291; Conservative 74; Mismatches 13; Indels 47; Gaps 11;

QY 2 LDIVWAAEAVPWSKTGGLGDTGGLPIELVKRGHRYMTIAPRYDQYADADMTSVVVDI- 60  
DB 87 MNLIEVGVTEAVPWSKTGGLGDTGGLPPLSANGHRVMIVPRYDQYADADMTNTIEVK 146

QY 61 ---MGEKVRYPHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117  
DB 147 VGDRTKVRFFCHFCRGVDFVDFHPIELFKVWGKTSKLYGPAAGDDYQDNOLRFSIFC 206

QY 118 KAAIEAARVLPF-----GP-GEDCVFAVDWHSALVPVLKDEYQPKGQTKAKSVLA 169  
DB 207 QAAIEAARVNLKSNKYFSGPYGDEVIFVANDWHSALFCYKMSYQSIGIFRNKVVFC 266

QY 170 IHNIAFOGRMWEAFKD---TKLPQAAFDKLFSDGYAKVYTEATPMEDEKPPITGKTY 226  
DB 267 IHNIAFOGRF---AFTDYSILNLPDQPKSSFDLDGHVK-----PIVG--- 306

QY 227 KKNLWKGIIAADKLVTSPNYATEIAADAAGGVELDTVIRAKIEGIVNGMDIEENP 286  
DB 307 KKNLWKGIIAADKLVTSPNYATEIAADAAGGVELDTVIRAKIEGIVNGMDIEENP 366

QY 287 KTDKFLSAPYDONSQVYAGKAAKALQALQELGLPVDPTAPLFAFGRLEEKGVDIILAAAL 346  
DB 367 STDKYISIKYDASTVLEKALKKEELQAEVCLPDKVNPVLIATFGRLEEKGSLLVEAI 426

QY 347 PKILATPKVQJAILGTCGAAYEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLP 406  
DB 427 PQFI-KENVQIVALTGKKEKEMKQLOLEISYDPKARGVAKFNVPVLAHMTAGADFIILIP 485

QY 407 SRPEPCGLIOLHAMHYGTVPVWASTGGVLDVTKGVTGFGHMGALNP--DKLDEADADALA 464  
DB 486 SRPEPCGLIOLQAMRYGTVPVWASTGGVLDVTKGVTGFGHMGALNP--DKLDEADADALA 545

QY 465 ATVRRASEVPAGGRYPVEMVANCISQDLSWSKPAQKWEGLL-----EEVYKGGVAT 516  
DB 546 KTVTKALGVYGTSAFAEMIKNCMAQELSWKPAKKEEVLNLCVDPSEFGIDGQ-BIAP 604

QY 517 AKKEIKVP 525  
DB 605 QAKENVATP 613

RESULT 3

Q9FR03 PRELIMINARY; PRT; 604 AA.

AC Q9FR03;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Granule-bound starch synthase.  
GN GBSSI.  
OS Perilla frutescens.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.  
OX NCBI\_TaxID=48386;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. OKDONG; TISSUE=DEVELOPING SEEDS;  
RA Hwang S.-K., Hwang Y.-S.;  
RT "Isolation and characterization of a cDNA encoding granule-bound  
starch synthase from perilla frutescens";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF210699; AAG43519.1;

DR	InterPro: IPR001296; Glycos_transf_1.	
DR	Pham: PF00534; Glycos_transf_1; 1.	
SQ	SEQUENCE 604 AA: 66214 MW: 435CE9601C85A5EF CRC64;	
	Query Match 43.0%; Score 1450.5; DB 10; Length 604;	
	Best Local Similarity 54.3%; Pred. No. 2.3e-77;	
	Matches 297; Conservative 62; Mismatches 145; Indels 43; Gaps 10;	
QY	1 2 LDIWMVAEAVPSKSTGGGLGVGTGGGLPIELVKGRHVTMTAPRYDQVADAWDTSVVYDI- 60	
DB	1 2 LDIWMVAEAVPSKSTGGGLGVGTGGGLPIELVKGRHVTMTAPRYDQVADAWDTSVVYDI- 60	
DB	78 MTLIFVAEAGCPWSKTGGLGVGVGGLPPLAANSHRVTMTSPRYDQVADAWDTSVVYDI 137	
QY	61 MGEKVR--YFHSIKKGVRWIDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117	
DB	61 MGEKVR--YFHSIKKGVRWIDHPFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117	
DB	138 VGDKEVTEGFHCYKRGVDRFVDHPFLFLEKVGWGTSKVYGPSGVYEDNQLRFSLLS 197	
QY	118 KAAIEAAARVLFF-----GP-GEDEVFVANDHSAIVPVLLKDEYQPKGQFTAKSVLA 169	
DB	118 KAAIEAAARVLFF-----GP-GEDEVFVANDHSAIVPVLLKDEYQPKGQFTAKSVLA 169	
DB	198 LAALEAPRVNLTSNKYFSGPYGEDVVFVANDHWTAVLPCLKTIYQPKGIYTNKAVVLC 257	
QY	170 IHNIAFOGRWEEAFKDTKLPOAFDKLAFSDGAKVYVTEATPMEEDEKPLTGTGYKKI 229	
DB	170 IHNIAFOGRWEEAFKDTKLPOAFDKLAFSDGAKVYVTEATPMEEDEKPLTGTGYKKI 229	
DB	258 IHNIAFOGRWEEAFKDTKLPOAFDKLAFSDGAKVYVTEATPMEEDEKPLTGTGYKKI 229	
QY	230 NNLKGGIIAADKLVTSFNAYATEAADAAGVGLDVIIRAKGIEGVNGMDIEEWNPKTD 289	
DB	230 NNLKGGIIAADKLVTSFNAYATEAADAAGVGLDVIIRAKGIEGVNGMDIEEWNPKTD 289	
DB	301 NNMKAGIIESRDLTVSPYANELVSGPDKGVELDNILRKCTVTGIVNGMDTQEWNPATD 360	
QY	290 KFLSAPYDONSYYAGKAAKALQALGLPVDPTAPLEAFICRLEEQGVDIILAALPKI 349	
DB	290 KFLSAPYDONSYYAGKAAKALQALGLPVDPTAPLEAFICRLEEQGVDIILAALPKI 349	
DB	361 KYIDNHVITITVMPDKPLLKALQAEVGLPDRNVPLVGFICRLEEQGSDIILVAALHKF 420	
QY	350 LATPKVQIAILGTGKAAVEKIVNAITGTYKGRAGVWKFSPAPLAHMLTAGADFMFLVPSRF 409	
DB	350 LATPKVQIAILGTGKAAVEKIVNAITGTYKGRAGVWKFSPAPLAHMLTAGADFMFLVPSRF 409	
DB	421 IEM-DVQVVILGTGKKEFEKQIEQLLELYPCKAVGVAKFNVPKHAITAGADFMFLVPSRF 479	
QY	410 EPCGLIIQLHAMHYGVTPVVAATGGGLVDVTYKGVGTGFHGMALNP--DKLDEADADALAATV 467	
DB	410 EPCGLIIQLHAMHYGVTPVVAATGGGLVDVTYKGVGTGFHGMALNP--DKLDEADADALAATV 467	
DB	480 EPCGLIIQLHAMHYGVTPICASTGGGLVDVTYKGVGTGFHGMALNP--DKLDEADADALAATV 539	
QY	468 BRASVEAGGRYPENMVANCIQSDLSWSKPAQKEGILLEVVYVGKG-----VATAK 518	
DB	468 BRASVEAGGRYPENMVANCIQSDLSWSKPAQKEGILLEVVYVGKG-----VATAK 518	
DB	540 GRALEVYGTAPAFREMINCMQLDLSWKGPAKWEIVL--LSLGVAGSEPGVGEDETAFLA 597	
QY	519 KEEIKVP 525	
DB	519 KEEIKVP 525	
DB	598 KENVATP 604	
DB	598 KENVATP 604	

```

RESULT 4
Q93VD9
ID Q93VD9 PRELIMINARY; PRT; 608 AA.
AC Q93VD9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Granule-bound starch synthase I (EC 2.4.1.21).
GN GBSSI.
OS Ipomoea batatas (Sweet potato) (Batate).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OC NCBI_TaxID=4120;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. KOREI 14;
RA Kimura T., Saito A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. KOREI 14;
RA Kimura T., Ideta O., Saito A.;
RT Identification of the gene encoding granule-bound starch synthase I

```

RT in sweet potato (*Ipomoea batatas* (L.) Lam.).";  
 RL Plant Biotechnol. 17:247-252(2000).  
 DR EMBL; AB071976; BAB68525.1; -;  
 DR EMBL; AB071604; BAB68126.1; -;  
 DR EMBL; AB071604; BAB68126.1; -;  
 DR InterPro: IPR001296; Glycos\_transf\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 DR Glycosyltransferase; transferase.  
 KW SEQUENCE 608 AA; 66689 MW; C93CDA49A1F50C03 CRC64;  
 SQ

Query Match	42.3%;	Score 1424.5;	DB 10;	Length 608;
Best Local Similarity	53.6%;	Pred. No. 7.9e-76;		
Matches 293;	Conservative 61;	Mismatches 150;	Indels 43;	Gaps 10;
Qy	2	LDIYVAAEAPW\$K\$TGGGLGVDVTGGPLIELVKGRHVRMTIAPRYDOYADAWDTSVVDI- 60		
Db	82	MNLVFGCEVGPWCKTGGGLGDLVGLP\$AARGHRVMTVCPRYDOYKADWDTCVVVELO 141		
Qy	61	MG----EKVRYPHSTKKGVRHVRWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHKRFALFC 117		
Db	142	VGDRIEPRFRF\$Y\$KRGYDRVDFVHPMFL\$K\$VWGKTSGLY\$PKAGKQYKDNQ\$RFL\$LC 201		
Qy	118	KAATEAA\$VLPF-----GP-GEDCVFVANDW\$HALVP\$LLKDBY\$QPKG\$OFT\$AK\$VLA 169		
Db	202	QAAL\$EAP\$V\$NLN\$SNSYFSGP\$YFVANDW\$HTALP\$CYL\$K\$MYQ\$RG\$Y\$MNAK\$VAF\$C 261		
Qy	170	IHN\$TAQ\$GRM\$EAF\$K\$D\$KLP\$QAAP\$DKL\$AF\$SD\$Y\$AK\$VY\$TEATP\$MEE\$DEK\$P\$LT\$GTY\$K\$K\$ 229		
Db	262	IHN\$TA\$YQ\$GR\$FA\$F\$D\$F\$S\$LLN\$L\$PDEY\$K\$S\$D\$F\$IDG\$Y\$K\$-----PV\$K\$G\$---R\$K\$ 304		
Qy	230	NWLKGGI\$TA\$AKL\$VTS\$PNYATE\$IA\$AD\$A\$AG\$V\$E\$LD\$TV\$IRAK\$IG\$IV\$NG\$MD\$TE\$EN\$PK\$D\$ 289		
Db	305	NW\$K\$AG\$IB\$EAD\$R\$V\$T\$V\$SP\$NYAK\$EL\$Y\$SC\$V\$K\$G\$Y\$E\$LD\$NH\$IRD\$CG\$IT\$G\$C\$NG\$MD\$TO\$E\$N\$PA\$ID\$ 364		
Qy	290	K\$LSAP\$Y\$D\$N\$S\$Y\$Y\$G\$K\$AA\$K\$Q\$AL\$Q\$E\$LG\$P\$D\$P\$TAP\$LF\$AF\$IG\$R\$E\$E\$Q\$K\$G\$V\$D\$I\$LAAL\$PK\$I\$ 349		
Db	365	K\$Y\$LV\$K\$Y\$D\$IT\$T\$Y\$W\$AK\$P\$LL\$K\$EAL\$Q\$A\$V\$GL\$P\$D\$EN\$P\$L\$F\$IG\$F\$IG\$R\$E\$E\$Q\$K\$D\$I\$Y\$A\$IS\$K\$F\$ 424		
Qy	350	LATP\$K\$VQ\$IA\$IL\$GT\$G\$K\$AA\$Y\$E\$K\$LY\$N\$A\$IG\$TK\$Y\$K\$R\$AK\$G\$V\$Y\$K\$E\$S\$AP\$LA\$H\$ML\$TAG\$AD\$F\$ML\$Y\$PS\$R\$F\$ 409		
Db	425	ISM-DVQ\$ILL\$G\$T\$K\$K\$K\$B\$Q\$E\$Q\$LE\$V\$M\$Y\$P\$D\$K\$ARG\$V\$AK\$EN\$V\$P\$LA\$H\$MT\$TAG\$AD\$F\$ML\$Y\$PS\$R\$F\$ 483		
Qy	410	EP\$C\$GL\$I\$QL\$H\$AM\$Y\$G\$IV\$P\$V\$V\$AST\$G\$GL\$Y\$D\$T\$Y\$K\$E\$G\$T\$G\$F\$H\$M\$G\$AL\$N\$P\$D\$---KL\$D\$E\$AD\$A\$LA\$A\$V\$T\$ 467		
Db	484	EP\$C\$GL\$I\$QL\$H\$AM\$Y\$G\$T\$P\$C\$A\$S\$T\$G\$GL\$Y\$D\$T\$Y\$K\$E\$G\$T\$G\$F\$H\$M\$G\$AF\$N\$V\$D\$C\$E\$T\$V\$D\$P\$ED\$V\$K\$V\$IT\$V\$T\$ 543		
Qy	468	RR\$A\$E\$V\$F\$AG\$G\$R\$Y\$P\$M\$V\$N\$A\$IS\$D\$LS\$W\$K\$S\$P\$AQ\$K\$W\$E\$G\$IL\$E\$E\$V\$Y\$Y\$K\$G\$K\$-----V\$AT\$AK\$ 518		
Db	544	GR\$AL\$AM\$Y\$G\$T\$LA\$T\$E\$M\$IK\$N\$C\$M\$S\$Q\$E\$LS\$W\$K\$G\$P\$AK\$N\$W\$E\$TV\$L\$--L\$S\$G\$V\$AG\$S\$E\$P\$G\$V\$E\$G\$E\$T\$AP\$LA\$ 601		
Qy	519	K\$E\$E\$K\$V\$P\$ 525		
Db	602	K\$EN\$V\$AT\$P\$ 608		

RESULT 5	
Q9XIS6	
ID	Q9XIS6 PRELIMINARY; PRT; 606 AA.
AC	C9XIS6;
DT	01-NOV-1999 (TrEMBLrel. 12, Created)
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Granule-bound starch synthase I precursor (EC 2.4.1.21).
GBSSI.	
GN	Phaseolus vulgaris (kidney bean) (French bean).
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX	NCBI_TaxID=3885;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Isono N., Nozaki K., Ito H., Matsui H., Honma M.;
RT	"Phaseolus vulgaris L. mRNA for Granule-Bound Starch Synthase I
RT	(GBSSI).";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR	EMBL; AB029546; BAA82346.1; -	
DR	InterPro; IPR001296; Glycos_transf.1.	
DR	Pfam; PF00534; Glycos_transf_1; 1.	
KW	Glycosyltransferase; Transferase; Transit peptide.	
FT	TRANSIT	78 POTENTIAL.
SC	SEQUENCE	506 AA: 67080 MW: 24604.3

Query Match 42.2%; Score 1423.5; DB 10; Le  
Best Local Similarity 51.5%; Pred. No. 9e-76;  
Matches 282; Conservative 82; Mismatches 120.

QY	2	LDIMVAAE	VAPWSKTGGLGDVNTGGLPIELVKGHRVMTIAPYDOYADAWDTSVVVDIM	61
Db		!!!	!!!	!!!
QY	80	NNLIFYCAE	VAPWSKTGGLGDVNLGSPSALAHEGHRVMTVSPYDOYKDAWDINVIVEVK	139
Db		!!!	!!!	!!!
QY	62	G----	EKVYFHSIKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHKRFALFC	117
Db		!!!	!!!	!!!
QY	140	VADRIETV	FFHCYKOGVDRVFDHPDFFLEKVGWKTGSKLYGFSAGVDYEDNQLRSLIC	199
Db		!!!	!!!	!!!
QY	118	KAATEARV	LPF-----GP-GEDEVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA	169
Db		!!!	!!!	!!!
QY	200	QAALEAPRV	LNLSNKNYFSSPYGEDVIFVANDWHTALLPCYLYKSMTQTRGVYRNTKVAF	259
Db		!!!	!!!	!!!
QY	170	IHNIAFOGR	WEEAFKDTKLP---QAADFKLAFSDGYAKVYTEATPMEEDKPPLTGKTY	226
Db		!!!	!!!	!!!
QY	260	IHNISYQGR	HPDEDFLLNLNPNYRSADF---FTDGHLK-----PVG-----	299
Db		!!!	!!!	!!!
QY	227	KKINMLKGI	IAADKLVTSPNKAETIADAGGVDELDTVIRAKGTIEGVINGMDTEWNP	286
Db		!!!	!!!	!!!
QY	300	KINMKWAA	ILSDVLVTSPYAKELVSGEDRGVELDNIIRKTGVAGIVNGMDIREWSP	359
Db		!!!	!!!	!!!
QY	287	KTDKFLSAP	YDONSVVYAGKAAAKEALQAEGLPVDPAPLFAFIGNLEQKQGVDIILAAL	346
Db		!!!	!!!	!!!
QY	360	KTDKFTID	FTDTSVKEAKFLKEALQAEVGLPVVNRDPLIGFIGLKEEQKSDILVEAI	419
Db		!!!	!!!	!!!
QY	347	PKILATPKV	QJAILGTGRAAYEKLVAIGTKYKGRAGVVKTSAPLAHMLTAGADFMVLP	406
Db		!!!	!!!	!!!
QY	420	PKFI--DQNV	QIIILGTGKSMKEQTEQLEEIYPEKARGIAKFDGGLPLAKHIIAGSFIMLP	478
Db		!!!	!!!	!!!
QY	407	SFEPCGLIOL	HAMHYGVVPPVASTGGVLDTVTKGVTFGPHMGALNPD--KLDDEADADALA	464
Db		!!!	!!!	!!!
QY	479	SFEPCGVLQ	HSMPYGVTPVVSSTGGVLDTVQEGFTGFHMGAFNWDCEAIDPADVEKTA	538
Db		!!!	!!!	!!!
QY	465	ATVRRASEV	FAGGRYPMMVANCISODLSWSKPAQKWEGLLERVYTKG-----GVATA	517
Db		!!!	!!!	!!!
QY	539	TVTRALTYGT	VAMEKIIQNCMAQDFSWKGPAKQWEKVLPSLDYGRSEAGIEGDEIAPL	598
Db		!!!	!!!	!!!
QY	518	KKEEIKVP	525	
Db		!!!	!!!	!!!
QY	599	AKENVATP	606	
Db		!!!	!!!	!!!

## RESULT 6

Q9ZSQ5	PRELIMINARY;	PRT;	607 AA.
Q9ZSQ5			
Q9ZSQ5			
AC			
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-MAR-2002	(TREMBlrel. 20, Last annotation update)	
DE	Granule-bound glycogen (starch) synthase (EC 2.4.1.11).		
DE	Astragalus membranaceus (Milk vetch) (Huang qi).		
SS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Astragalus.		
XX	NCBI_TaxID=83837;		
XX	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE-HAIRY ROOTS INDUCED BY RHIZOGENESIS AGROBACTERIUM LBA9402;		
RC	Peng J.S., Liu D., Hu Z.B.;		
CA	"CDNA cloning and structural analysis of granule-bound starch synthase		
VA	gene of Astragalus membranaceus."		
WT	Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.		

DR EMBL; AF097922; AAC70779.1; -.  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
DR KW Glycosyltransferase; Transferase.  
SQ SEQUENCE 607 AA; 66560 MW; 61BB0CACE94D387C CRC64;

Query Match 42.2%; Score 1423.5; DB 10; Length 607;  
Best Local Similarity 52.7%; Pred. No. 9.1e-76;  
Matches 289; Conservative 74; Mismatches 140; Indels 45.

QY	2	LDLVMVAEVAE	PWSVTGG	LGVDVTG	LPIELVK	RGRHVM	TIAPRYDQ	ADAWDT	SVVVDI	- 60	
Db	81	NNLVFEVGA	VPWSVTGG	LGVLG	PALAGN	HRVMTVS	PYDQ	KDAMD	IGVSVEIK	140	
QY	61	MG---	EKVRYFHS	IKKGVHR	WIDHP	FLAKVW	GKTSK	LYGPR	SAGADY	LDNKH	117
Db	141	VGDR	ETVTF	FFCYARG	VRDVF	DHPLF	LEKWK	GKTES	KLYGPK	TYGVYDK	100
QY	118	KAATEAAR	VLPP----	GP-	GEDCVF	VANDH	SHALV	PVLLK	DEYQ	KGQFTK	169
Db	201	QAAL	EAPRV	LNLSN	KNHFS	SGPYG	EDVVF	VANDH	WTALL	PCYLS	160
QY	170	IN	IAFOG	MEEA	FAKD---	TKL	QAPAD	KIAF	SDGYA	KVYTEAT	226
Db	261	IN	IAYOGR	---	HAF	SDLS	LNLSN	NEFR	SDF	IDGYD	300
QY	227	KK	INLKG	II	TAAD	KLVIV	SPNYA	TEATA	ADAG	VELD	286
Db	301	RK	INMKA	GVLES	ORVFT	VS	PYAKEL	VS	GEDR	GVELD	360
QY	287	KTDK	PLSAP	YQNS	YVAG	KAAK	EALQ	ELCP	VDPT	APLFA	346
Db	361	QD	TRID	VHYD	ASTV	TEAK	ILKEA	QAEV	GLP	VDNR	420
QY	347	PK	ILATP	KVQ	IA	LGTGA	AYE	KLVNA	IGTKY	GRAGV	406
Db	421	PK	FI-	DQNV	II	VLGTG	KKIME	QIE	LEV	YPGK	479
QY	407	SRE	PCGL	LIQ	LHAM	HYGT	VPV	ASTG	GLVT	VTEG	464
Db	480	SRE	PCGL	LIQ	LHAM	HYGT	VPV	ASTG	GLVT	VTEG	539
QY	465	AT	VRRAS	EV	FAGGR	YP	PWANC	TSOD	LSWS	KPAQ	517
Db	540	TV	NAL	KTYG	TQ	ALK	EMILN	CMQA	QDS	FWG	599
QY	518	KKEE	IKVP	525							
Db	600	AKEN	VATP	607							

## RESULT 7

	PRT;	608 AA.
QBYVUI	PRELIMINARY;	
IDAC	QBYVUI;	
AD	01-MAR-2002 (TREMblrel_20, Created)	
DT DT	01-MAR-2002 (T+EMBlrel_20, Last sequence update)	
EDIT	01-JUN-2002 (TREMblrel_21, Last annotation update)	
DDT	Granule binding starch synthase II.	
DE	Oryza sativa (Rice).	
OS	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
CC	Ehrhartoidae; Oryzaceae; Oryza.	
OC	NCL_TaxID=4530;	
XX	[1]	
RN	SEQUENCE FROM N.A.	
RP	Wu P., Dian W.M., Jiang H.W. ;	
RA	Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.	
AL	EMBL: AY069940; AAL59572.1; -.	
RR	InterPro: IPR001296; Glycos_transf_1.	
OR	Pham; PF00534; Glycos_transf_1; 1.	
SO	Sequence 608 AA; 67337 MW; D360662542E9A00E CRC64:	



Matches 277; Conservative 71; Mismatches 141; Indels 31; Gaps 7;	
QY	2 LDIYVMAAEVAPWSKTGGLGDTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVVDIM 61
DB	76 MNLVFGAEMAPWSKTGGLGDTGGLPMAANGHRVNVSPRIDQYKADAWDTSVISEIK 135
QY	62 ---GEKRYVPHSIKGVHRVWIDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
DB	136 VADEYERVFHCYKRGVDRVIDHPWLEKVRCKTEKIYGPDAITDYEDNQRFSLIC 195
QY	118 KAAIEAAKVL-----PF--GP-GEDEVFVANDHWSALVPLVLLKDEYKQGFTRAKSVLA 169
DB	196 QAAIEAPRILNINNPFESGYPGEDVVFVNDWHTGLLACLYLSNYQNGIYRTAKVAF 255
QY	170 IHNIAFOGMEWEEAFKDKLPQAFDKLAFSDGKAVKVTETATPMEDEKPPLTGTYKI 229
DB	256 IHNISYQGRSFDDFAQNLNLPDRFKSSFDIDGDK-----PVEG---RKI 298
QY	230 NWLKGIIIAADKLVTSPNYATEIAADAAGGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
DB	299 NWMKAGILQADKLVLSVPPYAEELISGEARGCELDNIMRLTGITGIVNGMDVSEWDPTKD 358
QY	290 KFLSAPYDONSIVAGKAAKAEALQAEGLGVPDTPAFIAGFGRLEQKGVDDIILALPKI 349
DB	359 KFLAVNYDITTALEAKALNKEALQAEVGLPDRKVPVAFIAGFGRLEQKGVDDIILALPKI 418
QY	350 LATPKVQITLGTGKAAVEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRF 409
DB	419 LKEEDVQIILLGTGKKEFKLKSMEEPGKRVAVRVPNAPLAHQMAGADLLAVTSRF 478
QY	410 EPCGLIQLHAMHYGTPVVPVASTGGLVDTYKGVTFGHMGALNPD--KLDEADADALAATV 467
DB	479 EPCGLIQLQMGYGPVPCVASTGGLVDTYKGVTFGHMGALNPD--KLDEADADALAATV 538
QY	468 RRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLLEEV 507
DB	539 KRAVKVGTGPAYQEMVKNCMIQDLSWKGPKNWEDVLEL 578
RESULT 10	
Q9MAQ0	PRELIMINARY; PRT; 610 AA.
ID	Q9MAQ0
AC	Q9MAQ0; 15, Created
DT	01-OCT-2000 (TREMREL. 15, Last sequence update)
DT	01-OCT-2000 (TREMREL. 15, Last sequence update)
DT	01-JUN-2001 (TREMREL. 17, Last annotation update)
DE	Granule-bound starch synthase.
FN	F9L11.8.
GN	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. COLOMBIA;
RA	Pederspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA	Altati H., Nguyen M., Lam B., Southwick A., Bei, Buehler E., Chin C.,
RA	Chiou J., Choi E., Dunn P., Gonzalez A., Howing B., Kim C., Koo T.,
RA	Lee J.M., Ienz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA	Pham P., Sakano H., Schwartz J., Shinn P., Thaveri A., Toriumi M.,
RA	Vaysberg M., Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AC006424; AAF31273.1;
DR	InterPro: IPR001296; Glycos.transf_1.
DR	Pham; PF00534; Glycos.transf_1; 1.
SQ	SEQUENCE 610 AA; 66879 MW; CF17F25BE12220DF CRC64;
Query Match 41.2%; Score 1390; DB 10; Length 610;	
Best Local Similarity 52.7%; Pred. No. 8.6e-74;	
Matches 287; Conservative 77; Mismatches 146; Indels 40; Gaps 12;	
QY	2 LDIYVMAAEVAPWSKTGGLGDTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVVDI- 60
Matches 277; Conservative 71; Mismatches 141; Indels 31; Gaps 7;	
QY	84 MSVFIGAEVGSWSTGGLGDTGGLPAPALARGHRVMTICPRYDOYKADAWDTCVVQIK 143
QY	61 MGEK---VRYPHSIKGVHRVWIDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117
DB	144 VGDKYENVRFFHCYKRGVDRVIDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 203
QY	118 KAAIEAAKVL-----GP-GEDEVFVANDHWSALVPLVLLKDEYKQGFTRAKSVLA 169
DB	204 QAAIEAPQVNLNNSKYSFSGYPGEDVVFVNDWHTALLPCYLSKYSQSGVYMNKRVFC 263
QY	170 IHNIAFOGMEWEEAFKDKLP---QAAEDKLAFLSDGKAVKVTETATPMEDEKPPLTGTY 226
DB	264 IHNIAFOGMEWEEAFKDKLP---QAAEDKLAFLSDGKAVKVTETATPMEDEKPPLTGTY 303
QY	227 KINNLKGGIIIAADKLVTSPNYATEIAADAAGGVVELDTVIRAKGIEGIVNGMDIEEWN 286
DB	304 KINNLKGGIIIAADKLVTSPNYATEIAADAAGGVVELDTVIRAKGIEGIVNGMDIEEWN 363
QY	287 KIDKFLSAPYDONSIVAGKAAKAEALQAEGLGVPDTPAFIAGFGRLEQKGVDDIILAL 346
DB	364 STDKYIDIKYDITTTVDKAPLKEALQAAVGLPDRVDFVIGFGRLEQKGVDDIILAL 423
QY	347 PKILATPKVQITLGTGKAAVEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMV 406
DB	424 SFPMGL-NVOMVILGTGKKEFKLKSMEEPGKRVAVRVPNAPLAHQMAGADLLAVTS 482
QY	407 SRFEPGGLIQLHAMHYGTPVVPVASTGGLVDTYKGVTFGHMGALNPD--DKLDEADADA 464
DB	483 SRFEPGGLIQLHAMHYGTPVVPVASTGGLVDTYKGVTFGHMGALNPD--DKLDEADADA 542
QY	465 AIVRRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVYVYGGVATAKKEIKV 524
DB	543 KAVTRAVAVGTSAMQEMVKNCMDQDFSWKGPAPLWKEKVLVLSLNV-A-GSEAGTEG 600
QY	525 PVAEK 529
DB	601 PLAKE 605
RESULT 11	
Q8S9C4	PRELIMINARY; PRT; 609 AA.
ID	Q8S9C4
AC	Q8S9C4; 21, Created
DT	01-JUN-2002 (TREMREL. 21, Last sequence update)
DT	01-JUN-2002 (TREMREL. 21, Last sequence update)
DT	01-JUN-2002 (TREMREL. 21, Last annotation update)
DE	Starch granule-bound starch synthase.
FN	WX-1.
GN	Oryza sativa (Rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehrhartideae; Oryzaceae; Oryza.
OX	NCBI_TaxID=4530;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. MILKY QUEEN;
RA	Sato H., Suzuki Y., Sakai M., Imbe T.;
RT	"Molecular Characterization of wx-1, a Novel Mutant Gene for Low-
RT	amylose Content in Endosperm of Rice (Oryza sativa L.).";
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AB066093; BAB88209.1;
SQ	SEQUENCE 609 AA; 66431 MW; B57CC13E0440D227 CRC64;
Query Match 41.2%; Score 1389.5; DB 10; Length 609;	
Best Local Similarity 51.2%; Pred. No. 9.2e-74;	
Matches 280; Conservative 76; Mismatches 148; Indels 43; Gaps 9;	
QY	2 LDIYVMAAEVAPWSKTGGLGDTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVVDI- 60
DB	83 MNVYVGAEMAPWSKTGGLGDTGGLPMAANGHRVNVISPRYDQYKADAWDTSVVAEIK 142
QY	61 ---MGEKRYVPHSIKGVHRVWIDHPFLAKVWGKTSKLYGPRSGADYLDNHRKRALFC 117



Db 143 VADRYVERFEHCYKHGVDRVDFIDHPSELEKVGWGTGKIYGDPTGVHDKDMQMFSLC 202  
 QY 118 KAAIDAAARVL-----PFGP--GDCVFNVDHWSALVPVLLKDEYQPKQGTAKSVLA 169  
 Db 203 QAALEAPRILNINNPFYKGYGDEWVFCNDWHTGPLASVLLKNNYQNGIYRNAKAVFC 262  
 QY 170 IHNIAFOGRMEEAFKDTKLQAAPDKLAFSDGYAKVYTEATPMEDEKPPLTGTYKKI 229  
 Db 263 IHNISYQGRFAFEDYPELNLSRFRSPDFIDGY-----DTPVEG-----RRI 305  
 QY 230 NMLKGGIITADKLVTVSPNYATEIAADAAGGVGLDVTIRAKGIEGIVNGMDIEENPKTD 289  
 Db 306 NNMKAGILEADRLTVSPVYAEELISGTARGCELDINMLTGITGIVNGMDVSEWDPSKD 365  
 QY 290 KFLSAPYDONSIVYAGAKAAKALQAEGLGVDPDTPAPLFAFTGRLEEQKGVDIILAAALPKI 349  
 Db 366 KYITAKYDATTAEAKALNKEALQAEGLGVDPDTPAPLFAFTGRLEEQKGVDPVMAAIPEL 425  
 QY 350 LATPKVQITAILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAFLAHLMTAGADFMLVPSRF 409  
 Db 426 M-QEDVQIVLLGTGKKKEKLLKSWEEKYPKVRVAVKFNAPLAHLIMAGADVLAVPSRF 484  
 QY 410 EPCGLIQLHAMHYGTPVPVASTGGVDTVKEGVTGFHMGALNPD--KLDEADADALAATV 467  
 Db 485 EPCGLIQLQGMRYGTPCACASTGGVDTVIEGKTGFHMGRLSDVCKVPESDVKVVAATL 544  
 QY 468 RRASEVFAAGGYPPEMVANCIQSOLSWSPKPAQKWEGLLEEVYVYGGK-----VATAK 518  
 Db 545 KRAKVGTGTAYEENVRNQMQLSWKGPKNWENVL--LGLGVAGSAPGIEGDELAPLA 602  
 QY 519 KEEIKVP 525  
 Db 603 KENVAAP 609

## RESULT 12

Q9SXK3 ID Q9SXK3 PRELIMINARY; PRT; 605 AA.  
 AC Q9SXK3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Starch synthase (GBSSI) (EC 2.4.1.21).  
 GN WAXY.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99321800; PubMed=10393240;  
 RA Murai J., Taira T., Ohta D.;  
 RT "Isolation and characterization of the three waxy genes encoding the  
 granule-bound starch synthase in hexaploid wheat."  
 RL Gene 234:71-79(1999).  
 DR EMBL; AB019623; BAA77351.1; --  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR InterPro; IPR000794; Ketoacyl-synt.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 605 AA; 66345 MW; 60A816276F78722B CRC64;

Query Match 41.0%; Score 1381; DB 10; Length 605;  
 Best Local Similarity 51.5%; Pred. No. 2.9e-73;  
 Matches 281; Conservative 72; Mismatches 153; Indels 40; Gaps 9;  
 QY 2 LDIIVMAAEVAPWSKTKGGLGDVTGGPLPIELVKKGRHVMYTIAPRYDQYADAWDTSVVVDIM 61  
 Db 78 MNLVFGAEMAPWSKTKGGLGDVTGGPLPIELVKKGRHVMYTIAPRYDQYADAWDTSVVSEIK 137

QY 62 ----GEKVYFHSIKKGVHRVWIDHPFLAKYVWGTGSKLYGPRSGADYLDNHRKRALFC 117  
 Db 138 VADRYVERFEHCYKHGVDRVDFIDHPSELEKVGWGTGKIYGDPTGVHDKDMQMFSLC 197  
 QY 118 KAAIDAAARVL-----PFGP--GDCVFNVDHWSALVPVLLKDEYQPKQGTAKSVLA 169  
 Db 203 QAALEAPRILNINNPFYKGYGDEWVFCNDWHTGPLASVLLKNNYQNGIYRNAKAVFC 262  
 QY 170 IHNIAFOGRMEEAFKDTKLQAAPDKLAFSDGYAKVYTEATPMEDEKPPLTGTYKKI 229  
 Db 263 IHNISYQGRFAFEDYPELNLSRFRSPDFIDGY-----DTPVEG-----RRI 305  
 QY 230 NMLKGGIITADKLVTVSPNYATEIAADAAGGVGLDVTIRAKGIEGIVNGMDIEENPKTD 289  
 Db 306 NNMKAGILEADRLTVSPVYAEELISGTARGCELDINMLTGITGIVNGMDVSEWDPSKD 365  
 QY 290 KFLSAPYDONSIVYAGAKAAKALQAEGLGVDPDTPAPLFAFTGRLEEQKGVDIILAAALPKI 349  
 Db 366 KYITAKYDATTAEAKALNKEALQAEGLGVDPDTPAPLFAFTGRLEEQKGVDPVMAAIPEL 425  
 QY 350 LATPKVQITAILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAFLAHLMTAGADFMLVPSRF 409  
 Db 426 M-QEDVQIVLLGTGKKKEKLLKSWEEKYPKVRVAVKFNAPLAHLIMAGADVLAVPSRF 484  
 QY 410 EPCGLIQLHAMHYGTPVPVASTGGVDTVKEGVTGFHMGALNPD--KLDEADADALAATV 467  
 Db 485 EPCGLIQLQGMRYGTPCACASTGGVDTVIEGKTGFHMGRLSDVCKVPESDVKVVAATL 544  
 QY 468 RRASEVFAAGGYPPEMVANCIQSOLSWSPKPAQKWEGLLEEVYVYGGK-----VYKGGVATAKK 519  
 Db 545 KRAKVGTGTAYEENVRNQMQLSWKGPKNWENVL--LGLGVAGSAPGIEGDELAPLA 599  
 QY 520 EEEKVP 525  
 Db 600 ENVAAP 605

## RESULT 13

Q9FUU6 ID Q9FUU6 PRELIMINARY; PRT; 605 AA.  
 AC Q9FUU6;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Granule bound starch synthase I (EC 2.4.1.21).  
 GN GBSSI.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. CHEYENNE; TISSUE=ENDOSPERM;  
 RA McCue K.F., Hurkman W.J., Tanaka C.K., Anderson O.D.;  
 RT "Starch Branching Enzymes Sbcl and Sbe2 from Wheat (Triticum aestivum  
 cv. Cheyenne): Molecular Characterization, Developmental Expression,  
 and Homolog Assignment by Differential PCR."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF286320; AAG27624.1; --  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR InterPro; IPR000794; Ketoacyl-synt.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 605 AA; 66326 MW; 21120D9D6F68B891 CRC64;

Query Match 40.9%; Score 1380; DB 10; Length 605;  
 Best Local Similarity 51.5%; Pred. No. 3.3e-73;  
 Matches 281; Conservative 72; Mismatches 153; Indels 40; Gaps 9;  
 QY 2 LDIIVMAAEVAPWSKTKGGLGDVTGGPLPIELVKKGRHVMYTIAPRYDQYADAWDTSVVVDIM 61  
 Db 78 MNLVFGAEMAPWSKTKGGLGDVTGGPLPIELVKKGRHVMYTIAPRYDQYADAWDTSVVSEIK 137

Db 78 NMLFVGVGEMAPWSKTGGGLDVLGGLPPAMAANGHRVMVISPRYDQYKDAWDTSVVSEIK 137  
 QY 62 ----GEKRYFHSIKGVRWIDHPWFLAKVWGKSGKLYGPRSCADYLDNHRKFAFC 117  
 Db 138 VADEYERVRFFHCYKRGVDRVFDHPCFFLEKVRGKTKKLYGPDAGTDYEDNQLRSLC 197  
 QY 118 KAAIEAARVLPF-----GP--GECVVFVANDWHSALVPLVLLKDEYQPKQGTAKASVLA 169  
 Db 198 QAALAPRILDLNNPIYFSGPYGVEDVFCNDHHTGLLACYLKSQSSGIYRTAKVAEC 257  
 QY 170 IHNIAPQGRMWEAFKDTKLPOAADFKLAFSDGYAKVYITATPMEDEKPLTKYKXI 229  
 Db 258 IHNISYQGRSFDFDQALNDRFKSSFDIDGYDK-----PVGG---RKI 300  
 QY 230 NMLKGGIIAADKLVTVPSPNYATEIAADAAGGVELDTVIRAKGIE-GIVNGMDIEEWNPKTD 289  
 Db 301 NNMKAGILOQADKLVITVSPYAEELISGEARGCELDNIMRLTGTGIVNGMDVSEWDPADK 360  
 QY 290 KFLSAPYDONSIVYAGKAAAEALQAEGLPVDPTAPLAFIFIGLEBKQGYDITLALPKI 349  
 Db 361 KFLAANYDYTTALEGKALNKEALQAEVGLPVDKVPFLVAFIGRLEBKQGYDVMIAAPEI 420  
 QY 350 LATPKVOIALGTGKAAAEKLVNAIGTKYKGRAGKGVVKSAPLAHMLTAGADFMVPSRF 409  
 Db 421 LKEEDVQIIVLGTGKKKKFELLKSVBEKFPKVRVVRFNAPLAHOMAGADVAVTSRF 480  
 QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDTVKEGVTGFHMGALNPD--KLDEADADALAATV 467  
 Db 481 EPCGLILOQGMRYGTPCACASTGGLVDTMEGKTGFHMGHLSVDCNVVEPADVKKVVTIL 540  
 QY 468 RRASEVFAAGRYPEMVCISQDLSWSKPAQKWEGLLEEV-----YVGKGVGATPAK 519  
 Db 541 KRAVKVGGTPAYHEMVKMNCMIQDLSNKGPAKNEDVLELGVGSEPGVIGE-EIAPLAM 599  
 QY 520 EETKVP 525  
 Db 600 ENVAAP 605

RESULT 14  
 Q9SQ58 PRELIMINARY; PRT; 599 AA.  
 ID Q9SQ58  
 AC Q9SQ58;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Granule-bound starch synthase GBSII.  
 OS Triticum aestivum (wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nakamura T., Vrinten P.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF109395; AAF14233.1;  
 DR InterPro: IPR001296; Glycos\_transf\_1.  
 DR Pfam: PF00534; Glycos\_transf\_1;  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 599 AA; 66045 MW; D486EFC90557F131 CRC64;

Query Match 40.4%; Score 1379; DB 10; Length 599;  
 Best Local Similarity 50.9%; Pred. No. 3.7e-73;  
 Matches 277; Conservative 78; Mismatches 147; Indels 48; Gaps 10;

QY 2 LDIVMAAEVAPWSKTGGGLDVLGGLPPAMAANGHRVMVISPRYDQYKDAWDTSVVVDIM 61  
 Db 72 MPIIFVTEVHPWCKTGGGLDVLGGLPPALAAAMGHRVMTAPRYDQYKDAWDTNVLVEVI 131  
 QY 62 ----GEKRYFHSIKGVRWIDHPWFLAKVWGKSGKLYGPRSCADYLDNHRKFAFC 117  
 Db 132 VGDRTVVRFFHCYKRGVDRVFDHPCFFLEKVRGKTKKLYGPDAGTDYEDNQLRSLC 191

QY 118 KAAIEAARVLPF-----GP--GECVVFVANDWHSALVPLVLLKDEYQPKQGTAKASVLA 169  
 Db 192 LAALAPRILDLNNPIYFSGPYGVEDVFCNDHHTGLLACYLKSQSSGIYRTAKVAEC 251  
 QY 170 IHNIAPQGRMWEAFKDTKLPOAADFKLAFSDGYAKVYITATPMEDEKPLTKYKXI 229  
 Db 252 IHNIAYQGRFPRVDFELLNLPESFMPSFDFVCGHVK-----PVVG---RKI 294  
 QY 230 NMLKGGIIAADKLVTVPSPNYATEIAADAAGGVELDTVIRAKGIE-GIVNGMDIEEWNPKT 288  
 Db 295 NNMKAGITCEDVVLITVSPHYVKELTSPEKGVGLDGLVIRAKPLETGTGIVNGMDVVDWNPAT 354  
 QY 289 DKFELSAPYDONSIVYAGKAAAEALQAEGLPVDPTAPLAFIFIGLEBKQGYDITLALPK 348  
 Db 355 DKYISYKYNATTVAERALKNEILQAEVGLPVDSSIPVIFIGRLEBKQGYDITLALPAIPE 414  
 QY 349 ILATPKVOIALGTGKAAAEKLVNAIGTKYKGRAGKGVVKSAPLAHMLTAGADFMVPSR 408  
 Db 415 FL--EENVQIIVLGTGKKKKFELLKSVBEKFPKVRVVRFNAPLAHOMAGADVAVTSRF 473  
 QY 409 FEPCLGTLQHAMHYGTVPVVASTGGLVDTVKEGVTGFHMGALNP--DKLDEADADALAAT 466  
 Db 474 FEPCLGTLQGMRYGTPICSTGGLVDTVSEGVTFHMGSFNVFETVDPADVAASVN 533  
 QY 467 VRRASEVFAAGRYPEMVCISQDLSWSKPAQKWEGLLEEVYKYG-----GVA 515  
 Db 534 VTRALKQYKTPSFHAMVQNCMAQDLSNKGPAKNW---EELLLGLGVGSEPGVIGE-EIAP 589  
 QY 516 TAKKEIKVP 525  
 Db 590 PLAKQNVATP 599

RESULT 15  
 Q43012 PRELIMINARY; PRT; 606 AA.  
 ID Q43012  
 AC Q43012;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Starch granule-bound starch synthase (EC 2.4.1.11).  
 GN WAXY.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SSP. L.F. SPONTANEA;  
 RA Wang Z.Y., Zheng F.Q., Gao J.P., Wang X.Q., Wu M., Zhang J.L.,  
 RA Hong M.M.;  
 RT "Identification of two transposon-like elements in rice wx gene."  
 RL Sci. China B 37:437-447(1994).  
 DR EMBL: X64108; CA445472.1;  
 DR InterPro: IPR001296; Glycos\_transf\_1.  
 DR InterPro: IPR000794; Ketoacyl-synt.  
 DR Pfam: PF00534; Glycos\_transf\_1;  
 DR PROSITE: PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 606 AA; 66323 MW; E1157468A5553CE CRC64;

Query Match 40.8%; Score 1376; DB 10; Length 606;  
 Best Local Similarity 50.7%; Pred. No. 5.7e-73;  
 Matches 277; Conservative 77; Mismatches 148; Indels 44; Gaps 10;

QY 2 LDIVMAAEVAPWSKTGGGLDVLGGLPPAMAANGHRVMVISPRYDQYKDAWDTSVVVDI- 60  
 Db 83 MNVFGEMAPWSKTGGGLDVLGGLPPAMAANGHRVMTAPRYDQYKDAWDTSVVAEIK 142  
 QY 61 ---MGEKRYFHSIKGVRWIDHPWFLAKVWGKSGKLYGPRSCADYLDNHRKFAFC 117  
 Db 143 VADRYERVRFFHCYKRGVDRVFDHPCFFLEKVRGKTKKLYGPDAGTDYEDNQLRSLC 202



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:49:26 ; Search time 11.5821 Seconds  
(without alignments)  
1653.790 Million cell updates/sec

Title: US-09-980-771a-3\_COPY\_58\_708

Perfect score: 3370

Sequence: 1 ALDINVVAEAFVPSWSTGGL.....SASKTSAKPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCUTS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1397.5	41.5	609	3	US-08-941-445A-7
2	1360.5	40.4	600	4	US-09-388-743-22
3	1353.5	40.2	533	3	US-08-941-445A-5
4	1342.5	39.8	616	4	US-09-388-743-14
5	1315	39.0	614	4	US-09-388-743-18
6	1314.5	39.0	615	4	US-09-388-743-2
7	902.5	26.8	641	4	US-08-836-567-10
8	882.5	26.2	649	4	US-09-192-909-2
9	881	26.1	671	4	US-09-196-390-2
10	873	25.9	801	4	US-09-388-743-26
11	871	25.8	767	4	US-08-836-567-8
12	861.5	25.6	539	3	US-08-941-445A-21
13	861.5	25.6	583	3	US-08-941-445A-13
14	858	25.5	690	4	US-09-388-743-6
15	845	25.1	558	4	US-08-836-567-6
16	841	25.0	799	4	US-09-196-390-6
17	837	24.8	698	3	US-08-941-445A-11
18	726	21.5	459	4	US-08-836-567-4
19	701.5	20.8	669	3	US-08-941-445A-9
20	580	17.2	477	1	US-07-735-065-2
21	580	17.2	477	1	US-08-469-202-12
22	580	17.2	477	2	US-08-484-434C-12
23	568.5	16.9	677	4	US-08-836-567-2
24	568.5	16.9	1197	4	US-08-836-567-12
25	568.5	16.9	1230	2	US-08-968-542C-35
26	545	16.2	735	4	US-09-115-704-2
27	530.5	15.7	533	4	US-09-388-743-10

28 504 15.0 1674 2 US-08-968-542C-12  
29 240 7.1 64 2 US-08-470-720-15  
30 215.5 6.4 79 2 US-08-470-720-13  
31 145.5 4.3 2035 1 US-08-046-585-5  
32 145.5 4.3 2035 1 US-08-393-703-5  
33 145.5 4.3 2035 5 PCT-US93-11721-5  
34 143.5 4.3 59 2 US-08-470-720-14  
35 134.5 4.0 1222 2 US-08-682-517-15  
36 134.5 4.0 1252 2 US-08-682-517-9  
37 132 3.9 520 4 US-09-000-016-7  
38 132 3.9 520 4 US-09-514-340-7  
39 132 3.9 734 4 US-09-000-016-4  
40 132 3.9 734 4 US-09-514-340-4  
41 132 3.9 823 4 US-09-000-016-2  
42 132 3.9 823 4 US-09-514-340-2  
43 129 3.8 30 2 US-08-470-720-9  
44 129 3.8 1125 4 US-09-513-783A-152  
45 129 3.8 1610 4 US-09-513-783A-22

#### ALIGNMENTS

RESULT 1  
US-08-941-445A-7  
; Sequence 7, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8089  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-941-445A-7

Query Match 41.5%; Score 1397.5; DB 3; Length 609;  
Best Local Similarity 51.4%; Pred. No. 1.4e-113;  
Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;  
QY 2 LDIVYAAEAVPWSKTGGLGVDVTGGPIELVKRGRHMTIAPRYDQYADAWDTSVVYDI- 60  
Db 83 MNVVFYGAEMAPWSKTGGLGVDVTGGPIELVKRGRHMTIAPRYDQYADAWDTSVVYAEIK 142









```
QY 3 DIVVAAEYAPWKTGGLGDTGGLPIELVKRHRVMTIAPRY-----DOYADAWDTS 55
Db 132 NIIFVTAEAAPYKTKGGLGDTGGLPVALAARGHRVMTIAPRY-----DOYADAWDTS 191
QY 56 V--VVDIMG--EKVRFHSHKGVHVRVMDHPFLAKVWCKTSGSKLYGPRSGADYLDNHK 111
Db 192 VRATVHCFGDAQVAYHEYRAGVDMVFDVHSSYC-----RPGTP-YGDIYGA-FGDNQF 244
QY 112 RFALFCKAAIEARVLPFGP--GEDCVFVANDHWSALVPVLLKDEYQPKQFTKAKSVL 168
Db 245 RFTLLSHAACEAPLVPLSGFTYGEKCLFLANDHWAALVPLLLAAKYRPGVYKARSIV 304
QY 169 AHNIAFQGRMWEAEKDKLPQAAFDKLPQAAFDKLPQAAFDKLPQAAFDKLPQAAFDK 228
Db 305 AHNIAHQQGVEPAVTVNLLGPPQWYGA---EWIFPTWARHALD-----TGET--- 351
QY 229 INWLGKGIITAAADKLVTVSPNYATPEIAADAAGGVDELTVI--RAKGIEGIVNGMDIEWNP 286
Db 352 VNVLKGAIAVADRILLVSGYSWEIIT-PEGGYGLHELSSROSIVNGITNGIDVNDWNP 410
QY 287 KTDKFLSAPYDQNSVYAKAAAEALQAEGLPVDPTAPLFAFIFGRLEQKGVYDIIILAA 346
Db 411 STDEHIAHYSINDL-SGKVOCKTDLQKELGLPIRPDCLPIGFIKGLDYQKGVYDIIILAI 469
QY 347 PKILATPKVQJAIILGTGKAAEKLVAIGTKYKRAKGVVYKTSAPLAHMLTAGADFMVLP 406
Db 470 PELMQN-DQVVMGSGEKOEDWMRHTENLFDKFRWVGVNVPVSHRITAGCDILLMP 528
QY 407 SRFPGLLIQLHAMHYGTVPVVASTGGLVDTVK-----EGV---TGFHMGALNPKDLD 456
Db 529 SRFPGLNQLYAMRYGTIPVHSTGGLRDTYKDFNPYAGGEGTGTFPSLTSEKL- 587
QY 457 EADADALATVRRASEVFAGGRYPE-----WVANCISODLSWSKPAQKWE 501
Db 588 ---LDTLKLAIAI-----GTYTEHKSSWEGLMRGMGRDYSWENAAIQYE 627

RESULT 8
US-09-192-909-2
; Sequence 2, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Froberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; TITLE OF INVENTION: starch synthases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
```

```
; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-192-909-2

Query Match 26.2%; Score 882.5; DB 4; Length 649;
Best Local Similarity 38.4%; Pred. No. 1.5e-69;
Matches 203; Conservative 95; Mismatches 174; Indels 57; Gaps 15;

QY 3 DIVVAAEYAPWKTGGLGDTGGLPIELVKRHRVMTIAPRY-----DOYADAWDTSV 56
Db 142 NIIFVTAEAAPYKTKGGLGDTGGLPVALAARGHRVMTIAPRY-----DOYADAWDTSV 201
QY 57 VVDIM--GE-KVRYFHSIKKGVHVRVMDHPFLAKVWCKTSGSKLYGPRSGADYLDNHK 112
Db 202 HIRIPCGGHEVTFHEYRDSVDVWFVDHPSY-----HRPGLNIGDKFEGA-FGDNQFR 254
QY 113 FALFCKAAIEARVLPFGP--GEDCVFVANDHWSALVPVLLKDEYQPKQFTKAKSVLA 169
Db 255 YTLCCYAAACEAPLVPLSGFTYGEKCLFLANDHWAALVPLLLAAKYRPGVYKARSIV 314
QY 170 IHNIAFQGRMWEAEKDKLPQAAFDKLPQAAFDKLPQAAFDKLPQAAFDKLPQAAFDK 229
Db 315 IHNIAHQQGVEPAVTVNLLGPPQWYGA---EWIFPTWARHALD-----TGET--- 361
QY 230 NWLGKGIITAAADKLVTVSPNYATPEIAADAAGGVDELTVI--RAKGIEGIVNGMDIEWNP 287
Db 362 NELKGAIVTADRIYTVSKGYSWEIIT-PEGGYGLHELSSROSIVNGITNGIDVNDWNP 420
QY 288 TDKFLSAPYDQNSVYAKAAAEALQAEGLPVDPTAPLFAFIFGRLEQKGVYDIIILAA 347
Db 421 TDKCIPCHYSVDDL-SGKAKCKGALQKELGLPIRPDCLPIGFIKGLDYQKGVYDIIILAI 479
QY 348 KILATPKVQJAIILGTGKAAEKLVAIGTKYKRAKGVVYKTSAPLAHMLTAGADFMVLP 407
Db 480 DLM-REDVQVMLGSGDPELEDWMRSTESIFDKFRGWRGVGVNVPVSHRITAGCDILLMP 538
QY 408 RFEPCGLLIQLHAMHYGTVPVVASTGGLVDTVK-----EGVTFGHMGALNPKDILDEA 458
Db 539 RFEPCGLNQLYAMRYGTIPVHSTGGLRDTYKDFNPYAGGEGTGTFPSLTSEKL- 594
QY 459 DADALATVRRASEVFAGGR--YPEMVANCISODLSWSKPAQKWEGLLE 505
Db 595 ---MLWTLKLAIAI-----GTYTEHKSSWEGLMRGMGRDYSWENAAIQYE 639

RESULT 9
US-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Lorz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
```





Db 51 IVFVIGESAPYAKSGGLGVDGSLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 110  
QY 58 VDIM---GE-KVRYFHSIKKGVHRVWIDHPWFLAKYWGKTSKLYGPRSGADYLDNHKRF 113  
Db 111 IRIPCFGGEHEVTFHEYRDSVDWVFDHPSY-----HRPGLNYGDKFKA-FGDNQFRI 163  
QY 114 ALFCAAIAEAARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAI 170  
Db 164 TLICAAACEAPLILELGGYIYGONCMFVNDWHSALVPVLLAARYPYGVYKDSRLVI 223  
QY 171 HNAIQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYVTEATMEDEKPLTGKTYKKIN 230  
Db 224 HNLHQGVPEASTYDGLGLPPEWYGALEW-----VPEWARHRLDKG-----EAVN 270  
QY 231 WLKGGIIAADKLVTVSPNYATEIAADAAGGVLDTVI--RAKIGIIVNGMDIEENPKT 288  
Db 271 FLKGAVVTADRIIVTSKYSWEVTT-AEGGQGLNELLSRKSVLNGVINGIDINDWNPAT 329  
QY 289 DKFLSAPYDONSVAAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAAALPK 348  
Db 330 DKCIPCHYSVDDL-SGKAKCKGALQKELGLPIRPDVLPIGFIGRLDYQKGDILQIIPD 388  
QY 349 ILATPKVQIATLGTGKAAEYKLVNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSR 408  
Db 389 LM-REDVQFVMLGSGDPELEDWMRSTESIFKDKFRGWGFSVPVSHRITAGCDILMPSR 447  
QY 409 FEPCGLIOLHAMHYGTVPVVASTGGVLDTVK-----EGVTGFHMGALNPKDL 455  
Db 448 FEPCGLNQLYAMQYGVTPVHATGGLRDIVENFPFGENGEGQGTGNAFAPLTIENM 503

## RESULT 13

US-08-941-445A-13  
; Sequence 13, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P.  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-941-445A-13

Query Match 25.6%; Score 861.5; DB 3; Length 583;  
Best Local Similarity 40.8%; Pred. No. 8.9e-67;  
Matches 194; Conservative 81; Mismatches 154; Indels 47; Gaps 13;  
QY 4 IVYMAAEVAPWKTGGLGVDVTGGLPIELVKRHRVMTIAPRY-----DOYADAWDTSVV 57  
Db 95 IVFVIGESAPYAKSGGLGVDGSLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 154  
QY 58 VDIM---GE-KVRYFHSIKKGVHRVWIDHPWFLAKYWGKTSKLYGPRSGADYLDNHKRF 113  
Db 155 IRIPCFGGEHEVTFHEYRDSVDWVFDHPSY-----HRPGLNYGDKFKA-FGDNQFRI 207  
QY 114 ALFCAAIAEAARVLPFGP---GEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAI 170  
Db 208 TLICAAACEAPLILELGGYIYGONCMFVNDWHSALVPVLLAARYPYGVYKDSRLVI 267  
QY 171 HNAIQGRMWEAFKDTKLPAQAFDKLAFSDGYAKVYVTEATMEDEKPLTGKTYKKIN 230  
Db 268 HNLHQGVPEASTYDGLGLPPEWYGALEW-----VPEWARHRLDKG-----EAVN 314  
QY 231 WLKGGIIAADKLVTVSPNYATEIAADAAGGVLDTVI--RAKIGIIVNGMDIEENPKT 288  
Db 315 FLKGAVVTADRIIVTSKYSWEVTT-AEGGQGLNELLSRKSVLNGVINGIDINDWNPAT 373  
QY 289 DKFLSAPYDONSVAAGKAAAEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAAALPK 348  
Db 374 DKCIPCHYSVDDL-SGKAKCKGALQKELGLPIRPDVLPIGFIGRLDYQKGDILQIIPD 432  
QY 349 ILATPKVQIATLGTGKAAEYKLVNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVPSR 408  
Db 433 LM-REDVQFVMLGSGDPELEDWMRSTESIFKDKFRGWGFSVPVSHRITAGCDILMPSR 491  
QY 409 FEPCGLIOLHAMHYGTVPVVASTGGVLDTVK-----EGVTGFHMGALNPKDL 455  
Db 492 FEPCGLNQLYAMQYGVTPVHATGGLRDIVENFPFGENGEGQGTGNAFAPLTIENM 547

## RESULT 14

US-09-388-743-6  
; Sequence 6, Application US/09388743  
; Patent No. 6423886  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Ian  
; TITLE OF INVENTION: No. 6423886el Starch Synthesis Polynucleotides and Their  
; TITLE OF INVENTION: Use in the Production of New Starches  
; FILE REFERENCE: 1144  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Curcuma zedoaria  
US-09-388-743-6

Query Match 25.5%; Score 858; DB 4; Length 690;  
Best Local Similarity 39.0%; Pred. No. 2.4e-66;  
Matches 208; Conservative 80; Mismatches 176; Indels 70; Gaps 15;  
QY 2 LDIVYMAAEVAPWKTGGLGVDVTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVV---- 57  
Db 199 MNILVAAECAPWKTGGLGVDVVGALPKALAKRHRVVMVSPRYGNVPPEKTEGNLKRKYK 258  
QY 58 VDIMGEKVRVYFHSIKKGVHRVWIDHPWFLAKYWGKTSKLYGPRSGADYLDNHKRFALFC 117  
Db 259 VDGQDMEIKYHYTIDSVDFVFDSP-----IFRHIGNDIY-----GNRVRDIILKRVLFC 309  
QY 118 KAAIEAARVLP-----FGPCEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLAIHN 172  
Db 310 KAAVEYPWHVPCGFCYGDG-NLVFIANDWHTSLLPVYLKACFDRDRGLMTYARCLLVIHN 368

QY 173 IAFQGRMWEAFKDTKLPOAFAFDKLAFLSDGYAKVYTEATPMEDEKPPITGKTYKKINWL 232  
Db 369 IAHQGRGLDDFSYVDLPHDHIDSPRLDD-----PVGGEHF-----NIF 407  
QY 233 KGGIIAADKLVTSVSPNYATEIAADAAGGVELDTVIRA--KGIEGVNGMDIEEWNPKTDK 290  
Db 408 AAGIARAARVTVSHGYAWELKT-SEGWGLHEIINECHKWPHGVINGIDITHSWPKEDA 466  
QY 291 FLSAP-----YDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIGRLEEQKGVDTIILAA 346  
Db 467 HLNSDGYNTLETLEMGKACQKAAALQREFGLPVRDDVPILLAFVIGRLDHQKGDILIAEAM 526  
QY 347 PKILATPKVOIALIGTGKAAAEKLYNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLP 406  
Db 527 -HWLWQDQLIMUGTGRPDLEMLRFEHREHGRKVGWGVFSVKAHHRITAGADALLMP 585  
QY 407 SRFEPCGLIQLHAMHYGTPVAVASTGGVLDVTKV-----GVTFGHMGALNPKDKLEADADA 462  
Db 586 SRFEPCGLNQLHAMHYGTPVAVHAYGGLRDTVQOQDFPNEIGLW-----TFDRAEAHR 639  
QY 463 LAA-----TVRRASEVFAGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVVY 509  
Db 640 MIVALGHCLNTYRNTKESWVGLQKRGMM-----QDLSWESAHEKVLVRAKY 688

## RESULT 15

US-08-836-567-6  
; Sequence 6, Application US/08836567  
; Patent No. 6130367  
; GENERAL INFORMATION:  
; APPLICANT: Kossmann, Jens  
; APPLICANT: Springer, Franziska  
; APPLICANT: Abel, Gernot  
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,567  
; FILING DATE: 24-JUL-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/04415  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 41 408.0  
; FILING DATE: 10-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: Agrevo-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 558 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-836-567-6

Query Match 25.1%; Score 845; DB 4; Length 558;  
Best Local Similarity 39.1%; Pred. No. 2.3e-65;  
Matches 207; Conservative 78; Mismatches 184; Indels 60; Gaps 14;  
QY 2 LDIVVAAEVAWPSWTTGGGLDVTGGTPIELVVRGHRVMTIAPRYDOYADAMTSTV-----V 57  
Db 67 KNIILVASCAPSWSTGGGLDVAGALPKALARHRVMTVAPRYDNPPEQDSGVKRIYK 126  
QY 58 VDIMGEKVRYPHSIKKGVHRVWDHPWFLAKVWGKTSKLYGPRSGADYLDHKKFALRC 117  
Db 127 VDGQDVEVIYFOAFIDGVDFVID-----SHMERHIGNNIY-----GGRNRVDILKRWLFEC 177  
QY 118 KAAIEAARVLP-----FGGEDCVFVANDWHVLPVLLKDEYQPKGQFTKAKSVLAHN 172  
Db 178 KAAIEVPWHPVPCGGVCYGDG-NLVFIANDWHHTALLPVYLKAYYRDNGIMNYTRSVLVIHN 236  
QY 173 IAFQGRMWEAFKDTKLPOAFAFDKLAFLSDGYAKVYTEATPMEDEKPPITGKTYKKINWL 232  
Db 237 IAHQGRGLDDFSYVDLPHDHIDSPRLDD-----PVGGEHF-----NIF 275  
QY 233 KGGIIAADKLVTSVSPNYATEIAADAAGGVELDTVIRAKG--JEGIVNGMDIEEWNPKTDK 290  
Db 276 AAGLKADRVTVSHGYSWELKT-SOGWGLHGIINENDWKLGIVNGIDTKEWNPDLV 334  
QY 291 FLSAP-----YDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFVIGRLEEQKGVDTIILAA 346  
Db 335 HLQSDGYMNSLDTLTGTGKPOCKAALQKELGLPVRDDVPILLAFVIGRLDHQKGDILIAEAS 394  
QY 347 PKILATPKVOIALIGTGKAAAEKLYNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLP 406  
Db 395 AMMG-ODVQLVMLGTGRDLEQMLRQFECQHNDKIRGWGVFSVKTSHRITAGADILLMP 453  
QY 407 SRFEPCGLIQLHAMHYGTPVAVASTGGVLDVTKV-----EGVTGFHMGALNPKDKLEADA 460  
Db 454 SRFEALRLNQLYAMKYGIPIVHAYGGLRDTVQOQDFPNEISGLWTFSPRAESQLIHAG 513  
QY 461 DALAATVRRASEVFAGRYPPEMVANCISQDLSWSKPAQKWEGLLEEVVY 509  
Db 514 NCL-LTYREYKKSWEG-----IQTRCMTQDLSWDNAQAQNYEEVLIAKY 556

Search completed: June 4, 2003, 14:58:35

Job time : 16.5821 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:56:21 ; Search time 20.1047 Seconds  
(without alignments)  
3342.965 Million cell updates/sec

Title: US-09-980-771a-3\_COPY\_58\_708  
Perfect score: 3370  
Sequence: 1 ALDIVMAAEVAPNSKTGGL.....SASKTAAKPLVSAATRKSA 651

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450.5	43.0	636	9 US-10-138-075-4	Sequence 4, Appli
2	1433.5	42.5	609	9 US-10-138-075-2	Sequence 2, Appli
3	1379	40.9	599	9 US-10-138-075-5	Sequence 5, Appli
4	881	26.1	671	10 US-09-952-677-2	Sequence 2, Appli
5	841	25.0	799	10 US-09-952-677-6	Sequence 6, Appli
6	555	16.5	915	9 US-10-163-214-6	Sequence 6, Appli
7	549	16.3	914	9 US-10-163-214-12	Sequence 12, Appli
8	545	16.2	909	9 US-10-163-214-2	Sequence 2, Appli
9	537	15.9	874	9 US-10-163-214-13	Sequence 13, Appli
10	410.5	12.2	474	12 US-10-007-693-99	Sequence 99, Appli
11	402.5	11.9	476	12 US-10-007-693-69	Sequence 69, Appli
12	385.5	11.4	293	9 US-10-163-214-10	Sequence 10, Appli
13	215.5	6.4	117	10 US-09-739-438-2	Sequence 2, Appli
14	206	6.1	409	9 US-09-738-626-4738	Sequence 4738, Ap
15	189	5.6	381	9 US-09-738-626-5896	Sequence 5896, Ap
16	164	4.9	143	10 US-09-739-438-4	Sequence 4, Appli
17	155.5	4.6	191	10 US-09-924-358-22	Sequence 22, Appli
18	155.5	4.6	418	9 US-09-738-626-3951	Sequence 3951, Ap
19	136.5	4.1	385	10 US-09-767-041-18	Sequence 18, Appli

20	135	4.0	2993	9 US-09-738-626-6239	Sequence 6239, Ap
21	134.5	4.0	1222	10 US-09-137-531-15	Sequence 15, Appli
22	134.5	4.0	1252	10 US-09-137-531-9	Sequence 9, Appli
23	129	3.8	1125	9 US-10-100-957A-152	Sequence 152, App
24	129	3.8	1610	9 US-10-100-957A-22	Sequence 22, Appli
25	128	3.8	2910	9 US-10-124-800-2	Sequence 2, Appli
26	126	3.7	387	9 US-09-738-626-3890	Sequence 3890, Ap
27	125.5	3.7	398	10 US-09-934-899-16	Sequence 16, Appli
28	125.5	3.7	398	10 US-09-934-899-36	Sequence 36, Appli
29	124.5	3.7	1704	9 US-09-991-262-40	Sequence 40, Appli
30	124	3.7	379	9 US-09-738-626-6655	Sequence 6655, Ap
31	124	3.7	1461	9 US-10-021-955-88	Sequence 88, Appli
32	123.5	3.7	413	9 US-09-738-626-4025	Sequence 4025, Ap
33	123	3.6	1528	9 US-09-945-917-3	Sequence 3, Appli
34	123	3.6	1583	9 US-09-945-917-4	Sequence 4, Appli
35	121	3.6	339	9 US-09-847-208-116	Sequence 116, App
36	118.5	3.5	372	10 US-09-815-242-5277	Sequence 5277, Ap
37	118.5	3.5	384	10 US-09-815-242-12605	Sequence 12605, A
38	118	3.5	267	10 US-09-852-555-2	Sequence 2, Appli
39	118	3.5	333	9 US-09-847-208-147	Sequence 147, App
40	118	3.5	1463	9 US-09-971-536-69	Sequence 69, Appli
41	117.5	3.5	636	9 US-09-738-626-5590	Sequence 5590, Ap
42	117	3.5	3241	10 US-09-841-786-1	Sequence 1, Appli
43	117	3.5	3739	9 US-09-860-846-33	Sequence 33, Appli
44	117	3.5	3739	9 US-09-988-384B-33	Sequence 33, Appli
45	117	3.5	3739	9 US-09-836-821-33	Sequence 33, Appli

## ALIGNMENTS

## RESULT 1

US-10-138-075-4  
; Sequence 4, Application US/10138075  
; Publication No. US20030087369A1  
; GENERAL INFORMATION:  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Orozco, Emil M.  
; TITLE OF INVENTION: Granule-Bound Starch Synthase  
; FILE REFERENCE: BB1474 NA  
; CURRENT APPLICATION NUMBER: US/10/138,075  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/288,315  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-138-075-4

Query Match	43.0%;	Score 1450.5;	DB 9;	Length 636;
Best Local Similarity	53.0%;	Pred. No. 6.2e-95;		
Matches 291;	Conservative 71;	Mismatches 144;	Indels 43;	Gaps 10;
QY	2	LDIVMAAEVAPNSKTGGLGDDVTGGTPIELVGRHVRVMTIAPRYDQYADAWDTSVVVDI-	60	
Db	106	MFFIIIGTEVPWCVTGGGLGDLGPPALAGFGRHVRVMTIAPRYDQYKDAWDTSVVIVK	165	
QY	61	---MGEKRYFHSIKKGVHRVWIDHPWFIKAVMGTKSGKLYGPRSGADYLDNHRKFLFC	117	
Db	166	VCDREKVRFFHCYKRGVDRVFDHPWFLKVGKTKGKLYGPTTGNEDYNQLRSLFC	225	
QY	118	KAATEAARVLPF-----GP-GEDEVFANDWHSALVPVLKDFYQPKQFTKAKSVLA	169	
Db	226	QAALAPRVLSLNSKSYFSGPYGDEVIFVANDWHTALIPCYLKSMSYQSGIYNARVFC	285	
QY	170	IHTIAFGRMWEAEAFKDKLPQAAFDKLAFSQYAKVYTEATPMEDEKPPITGTYKKI	229	

Db 286 IHNTAYOGRFAFADFSLNLPDQFKSSPDFIDGHVK-----PVVG---RKI 328  
QY 230 NWLKGIIAADKLVTSPNYATEAADAAGGVELDTVIRA-----KGIEGVNGMDIEWN 285  
Db 329 NWLKGIIAADKLVTSPNYATEAADAAGGVELDTVIRA-----KGIEGVNGMDIEWN 388  
QY 286 PKTDFLSAPYDONSUYAGKAAKAEALQAEGLPVDPTAPLFAFAGTGRLEEQKGVDIILAA 345  
Db 389 PTTDKYIAVYDVSIVLEAKALLKEALQAEGLPVDNRNPLIFIGTGRLEEQKGVDIILAA 448  
QY 346 LPKILATPKVQIAIILGTGKAAAYEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMV 405  
Db 449 IPQFT-KONVOLVALGTGKOMEKOLELETSYDPKARGVAKFNVPFLAHMIIAGADFILV 507  
QY 406 PSRFPCCGLIOLHMHYGTVPVASTGGGLVDVFKGVTFGFMALNP--DKLDEADADAAL 463  
Db 508 PSRFPCCGLIOLHMHYGTVPVASTGGGLVDVFKGVTFGFMALNP--DKLDEADADAAL 567  
QY 464 AATVRRASEVAGGAGGYPENAVNCISQDLSWSKPAKWEGLLEEVYVY-----GKGG--VAT 516  
Db 568 SKTVRALAVYGTAFTEIKNMAQDLSWKPAKWEGLLEEVYVY-----GKGG--VAT 627  
QY 517 AKKEEIKVP 525  
Db 628 QAKENVATP 636  
RESULT 2  
US-10-138-075-2  
; Sequence 2, Application US/10138075  
; Publication No. US20030087369A1  
; GENERAL INFORMATION:  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Orozco, Emil M.  
; TITLE OF INVENTION: Granule-Bound Starch Synthase  
; FILE REFERENCE: BB1474 NA  
; CURRENT APPLICATION NUMBER: US/10/138,075  
; PRIOR FILING DATE: 2002-05-02  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 609  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-138-075-2

Query Match 42.5%; Score 1433.5; DB 9; Length 609;  
Best Local Similarity 52.3%; Pred. No. 9.4e-94;  
Matches 288; Conservative 75; Mismatches 139; Indels 49; Gaps 11;  
QY 2 LDIVMVAEAPWPKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTSSVVDIM 60  
Db 81 MTIVFVATEVHPWCKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTSSVVDIM 140  
QY 61 MG---EKVYFHSIKKGVHRVWIDHPWFLAKYWGKGTGSKLYGPRSGADYLDNKKRALFC 117  
Db 141 IGDVTETVFFHCYKRGVDRVFDHPWFLAKYWGKGTGSKLYGPRSGADYLDNKKRALFC 200  
QY 118 KAAIEAARVLPF-----GP-CEDCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169  
Db 201 LAALAPRVNLNNSSEYFSGPYGVNFFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 260  
QY 170 IHNTAYOGRFAFADFSLNLPDQFKSSPDFIDGHVK-----PVVG---RKI 229  
Db 261 IHNTAYOGRFAFADFSLNLPDQFKSSPDFIDGHVK-----PVVG---RKI 303  
QY 230 NWLKGIIAADKLVTSPNYATEAADAAGGVELDTVIRA-----KGIEGVNGMDIEWN 288

Db 304 NNMKAGITIEDLVLTVSPHYVKELTSGDPKGVDELGVLRKPLEIGIVNGMDVYEDPST 363  
QY 289 DKFLSAPYDONSUYAGKAAKAEALQAEGLPVDPTAPLFAFAGTGRLEEQKGVDIILAA 348  
Db 364 DKFLSAPYDONSUYAGKAAKAEALQAEGLPVDPTAPLFAFAGTGRLEEQKGVDIILAA 423  
QY 349 ILATPKVQIAIILGTGKAAAYEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMV 408  
Db 424 FVG-ENVQIIVLGTGKMEELTQLEVKYPPNARGIAKFNVPFLAHMIFAGADFIIVPSR 482  
QY 409 FPGCGLIOLHMHYGTVPVASTGGGLVDVFKGVTFGFMALNP--DKLDEADADAAL 466  
Db 483 FPGCGLIOLHMHYGTVPVASTGGGLVDVFKGVTFGFMALNP--DKLDEADADAAL 542  
QY 467 VRRASEVAGGAGGYPENAVNCISQDLSWSKPAKWEGLLEEVYVY-----GKGG--VAT 514  
Db 543 VTRALKOYDTPAFHEMVQNCMAQDLSWKPAKWEGLLEEVYVY-----GKGG--VAT 598  
QY 515 ATAKEEIKVP 525  
Db 599 APLAKENVATP 609  
RESULT 3  
US-10-138-075-5  
; Sequence 5, Application US/10138075  
; Publication No. US20030087369A1  
; GENERAL INFORMATION:  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Orozco, Emil M.  
; TITLE OF INVENTION: Granule-Bound Starch Synthase  
; FILE REFERENCE: BB1474 NA  
; CURRENT APPLICATION NUMBER: US/10/138,075  
; PRIOR FILING DATE: 2002-05-02  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-138-075-5

Query Match 40.9%; Score 1379; DB 9; Length 599;  
Best Local Similarity 50.4%; Pred. No. 6.8e-90;  
Matches 277; Conservative 78; Mismatches 147; Indels 48; Gaps 10;  
QY 2 LDIVMVAEAPWPKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTSSVVDIM 61  
Db 72 MPILFVATEVHPWCKTGGGLDVTGGLPIELVKRGHRVMTIAPRYDOYADAWDTSSVVDIM 131  
QY 62 ---GKVFYFHSIKKGVHRVWIDHPWFLAKYWGKGTGSKLYGPRSGADYLDNKKRALFC 117  
Db 132 VGRDTEVTRFFHCYKRGVDRVFDHPWFLAKYWGKGTGSKLYGPRSGADYLDNKKRALFC 191  
QY 118 KAAIEAARVLPF-----GP-CEDCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169  
Db 192 LAALAPRVNLNNSSEYFSGPYGVNFFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 251  
QY 170 IHNTAYOGRFAFADFSLNLPDQFKSSPDFIDGHVK-----PVVG---RKI 229  
Db 252 IHNTAYOGRFAFADFSLNLPDQFKSSPDFIDGHVK-----PVVG---RKI 294  
QY 230 NWLKGIIAADKLVTSPNYATEAADAAGGVELDTVIRA-----KGIEGVNGMDIEWN 288  
Db 295 NNMKAGITIEDLVLTVSPHYVKELTSGDPKGVDELGVLRKPLEIGIVNGMDVYEDPST 363  
QY 289 DKFLSAPYDONSUYAGKAAKAEALQAEGLPVDPTAPLFAFAGTGRLEEQKGVDIILAA 348



Db 355 DKVISYKYNATTVAERALKNEILQAEVGLPVDSSIPVIVFGRLEQKSGDILIAIPE 414  
QY 349 ILATPKVQIALGTGKAAEKIUNATGTYKGRAGKGVVKSAPLAHMLTAGADFMVPSR 408  
Db 415 FL-EENVQIIVLGTGKKMEELMELEAKYPQNGARTAKENVPLAHMFPAGANFIIVPSR 473  
QY 409 FEPGGLIOLHAMHYGTVPVVASTGGGLVDVTKEGVTGFHMGALNP--DKLDEADADALAT 466  
Db 474 FEPGGLIOLQGMRYGIVPICSSTGGGLVDVTSEGVTFGHMGSEFVFETVDPADVAASN 533  
QY 467 VRRASEVFFAGGRYPENVANCISODLSWSKPAQKWEGLLEEVYKKG-----GVA 515  
Db 534 VTRALQXKTPPSHANVQNCMAQDLSWKGPAKW-----EALILGLVGSQPGIEGEIA 589  
QY 516 TAKKEIKVP 525  
Db 590 PLAKQNVATP 599

RESULT 4  
US-09-952-677-2  
; Sequence 2, Application US/09952677  
; Patent No. US20020138876A1  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Luticke, Stephanie  
; Walter, Lennart  
; Froberg, Claus  
; Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/952,677  
; FILING DATE: 14-Sep-2001

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,390  
; FILING DATE: 19-NOV-2002  
; APPLICATION NUMBER: DE 196 21 588.9  
; FILING DATE: 29-MAY-1996  
; APPLICATION NUMBER: DE 196 36 917.7  
; FILING DATE: 11-SEP-1996  
; APPLICATION NUMBER: PCT/EP97/02793  
; FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:  
; NAME: Haley, Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: AGREVO-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 671 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-952-677-2

Query Match 26.1%; Score 881; DB 10; Length 671;  
Best Local Similarity 35.2%; Pred. No. 1.9e-54;  
Matches 237; Conservative 94; Mismatches 225; Indels 118; Gaps 22;  
QY 4 IVWVAARVAPWSTGGGLGDTVTGGLPIELVGRHVRMTIAPRY-----DOYADAWDTSVV 57  
Db 56 IVEVTGTAAPYAKSGGLGVCGLPIALAAAGHRVVMVMPRYLNGSSDKNYAKALYAKH 115  
QY 58 VDIM----GEKVYFEHSIKKGVRHVDHDFWELAKVWCKSKLYGPRSGADYLDNKRKF 113  
Db 116 IKTPCGSGSEVTFHEYRONVDMVFDHP-----SYHREPS-LYDNFNGA-FGDNQF 168  
QY 114 ALFCKAAIEAARVLPFGP---GEDCVFVANDWHSALVPVLLKDYQPKGQFTAKSVIAI 170  
Db 169 TLLCYAAACEAPLILELGGYIYGQNCMFVVDWHSALVPVLLAAKYRYGVYVRSRSLVI 228  
QY 171 HNTAFQGRWEEAFKDKLPQAAFDKLAFSDDGAKVYVTEATPMPEDEKPLTGKTYKKN 230  
Db 229 HNLHQGVPEASTYDGLPPEWYGALEW-----VPEWARHAKDKG-----EAVN 275  
QY 231 WLKGGIITADKLVTVSPNYATEIAADAAGGVDELDTVI--RAKGIETGVNMDIEENPKT 288  
Db 276 FLKGVNADRIYVTVSGYSEWYTT-AEGGQGLNELLSRSKSVLNGVINGIDINDWPTT 334  
QY 289 DKLSAPYDQNSYAGKAAKAEALQAEGLPVDPTAFPLFAGFGRLEEKQGVDIILALPK 348  
Db 335 DKCLPHYSVDDL-SGKAKCKAELOKELGLPVREDVPLIGFGRLDYQKGLDKLKAPE 393  
QY 349 ILATPKVOIALGTGKAAEKIUNATGTYKGRAGKGVVKSAPLAHMLTAGADFMVPSR 408  
Db 394 LM-REDVQVYMLGSGDPIFEGWNRSTESSYKDKFERGVWGSFVPSVSHRITAGCDILLMP 452  
QY 409 FEPGGLIOLHAMHYGTVPVVASTGGGLVDTVK-----EGVTGPHMGALNPKDLDEAD 459  
Db 453 FEPGGLNQLXAMQYGVVPHVGGTGLDVTETNPFCAKGEETGWAFSPLTVDKMLWA- 511  
QY 460 ADALATVRRASEVFAGGRYPENVANCISODLSWSKPAQKWEGLLEEVYKKGVAFAK 519  
Db 512 -----LRTAMSTFE-----HKPS--WEGLM-----KR 532  
QY 520 EEIKVPVPAEKIPGDLPAVSYAPNTLKPVSASVEGNGAAAPKVGTTPAMGAWRATTPSGP 579  
Db 533 GMTKHTWDHAPSTSRSSSGPSWTNPTSC-----RGLGRSKCESPSAL 577  
QY 580 SPAAT---PKVTYTPALPATAKPKTAGLK-LAGEASTTSTSENGAASNGNGASASK 635  
Db 578 KTSSSFRGPEGYPCILRCPATVESQCACLLWFAG-----SRTYDCCA-----AAAVT 625  
QY 636 TSAAKPLVSAATRK 649  
Db 626 ASGGRQLQFWGIRK 639

RESULT 5  
US-09-952-677-6  
; Sequence 6, Application US/09952677  
; Patent No. US20020138876A1  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Luticke, Stephanie  
; Walter, Lennart  
; Froberg, Claus  
; Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York

STATE: New York  
COUNTRY: United States of America  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/952,677  
FILING DATE: 14-Sep-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/196,390  
FILING DATE: 19-NOV-2002/38876A1-1998  
APPLICATION NUMBER: DE 196 21 588.9  
FILING DATE: 29-MAY-1996  
APPLICATION NUMBER: DE 196 36 917.7  
FILING DATE: 11-SEP-1996  
APPLICATION NUMBER: PCT/EP97/02793  
FILING DATE: 28-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 799 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-952-677-6

Query Match 25.0%; Score 841; DB 10; Length 799;  
Best Local Similarity 39.1%; Pred. No. 1.7e-51;  
Matches 208; Conservative 75; Mismatches 183; Indels 66; Gaps 15;

QY 2 LDIVVAAEAPVAPWSKTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWTSV--VVD 59  
Db MNVVAECSWCKTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWTSV--VVD 59  
QY 60 IMGE--KVRYFHSIKKGVHRVWIDHPFLAK--VWGKTGSKLYGPRSGADYLDNHKREA 114  
Db AAGQDMVNYFAYIDGVDFVDAFLPHRQEDYIGGSRQELM-----KRM 415  
QY 115 LFCKAAIEA-----ARVLPGGDCVFNVDHWSALVPLVKDEYQPKQGTAKSVLA 169  
Db LFCKAAIEA-----ARVLPGGDCVFNVDHWSALVPLVKDEYQPKQGTAKSVLA 169  
QY 416 LFCKAAIEA-----ARVLPGGDCVFNVDHWSALVPLVKDEYQPKQGTAKSVLA 169  
Db LFCKAAIEA-----ARVLPGGDCVFNVDHWSALVPLVKDEYQPKQGTAKSVLA 169  
QY 170 INIAFGGRMWEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGKTKYKI 229  
Db INIAFGGRMWEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGKTKYKI 229  
QY 475 INIAFGGRMWEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGKTKYKI 229  
Db INIAFGGRMWEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGKTKYKI 229  
QY 230 NWLKGITAAADKLVTSPNYATEIAADAGAGVGLDVTIRAKG--IEGIYNGMDIEFWNPK 287  
Db NWLKGITAAADKLVTSPNYATEIAADAGAGVGLDVTIRAKG--IEGIYNGMDIEFWNPK 287  
QY 514 NYFAAGLKMAQGVVWVSPGLYELWELKT--VEGGWGLDIIHQNDWKTRGIVNGIDNMEWNP 572  
Db NYFAAGLKMAQGVVWVSPGLYELWELKT--VEGGWGLDIIHQNDWKTRGIVNGIDNMEWNP 572  
QY 288 TDKFTLSAP----YDQNSVYAGAAKAEALQAEIGL--PVDPAPLFAFIEGKQGVDIIL 343  
Db TDKFTLSAP----YDQNSVYAGAAKAEALQAEIGL--PVDPAPLFAFIEGKQGVDIIL 343  
QY 573 VDAHLKSDGYTWFSLDTSGRKQCKEALQRELQVLRADVPLFLGFLDQKGVVETIA 632  
Db VDAHLKSDGYTWFSLDTSGRKQCKEALQRELQVLRADVPLFLGFLDQKGVVETIA 632  
QY 344 AALPKLAPKPVQIATIGTGAAYEKLVAIGTKYKRAKGVVVKFSAPLAHMLTAGADFW 403  
Db AALPKLAPKPVQIATIGTGAAYEKLVAIGTKYKRAKGVVVKFSAPLAHMLTAGADFW 403  
QY 633 DAMPRIV--SQDQVLMGLGRHDLSEMLQHFREHDKVRGVNGVSVRLAHLITAGADAL 691  
Db DAMPRIV--SQDQVLMGLGRHDLSEMLQHFREHDKVRGVNGVSVRLAHLITAGADAL 691  
QY 404 LVPSRPEPCGLIQLHAMHYGTVPVAVSTGGVLDTV-----KEGVTGFHMGALNPDKLDE 457  
Db LVPSRPEPCGLIQLHAMHYGTVPVAVSTGGVLDTV-----KEGVTGFHMGALNPDKLDE 457  
QY 692 LMPSPFPCGLINQLAMAYCTVPVAVHVGGLRDTVPVPPFNHSGGLGWTDFDRAEAKHLE 751  
Db LMPSPFPCGLINQLAMAYCTVPVAVHVGGLRDTVPVPPFNHSGGLGWTDFDRAEAKHLE 751  
QY 458 ADADALAAVTRASEVFAAGRYPEMVAVNCISQDLSWSKFAQKWEGLLEEVV 509

Db 752 ALGHCL-RTYRDFKE-----SWRALQERQMSQDTSWEHAAKLYEDVLVRKY 797  
RESULT 6  
US-10-163-214-6  
; Sequence 6, Application US/10163214  
; Publication NO. US20030097688A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen M.  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Thorpe, Catherine J.  
; TITLE OF INVENTION: Starch Synthase Isoform V  
; FILE REFERENCE: BB1520 US NA  
; CURRENT APPLICATION NUMBER: US/10/163,214  
; CURRENT FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/297,099  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 915  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-163-214-6

Query Match 16.5%; Score 555; DB 9; Length 915;  
Best Local Similarity 29.5%; Pred. No. 4.2e-31;  
Matches 164; Conservative 92; Mismatches 183; Indels 116; Gaps 20;  
QY 2 LDIVVAAEAPVAPWSKTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWTSV--VVD 60  
Db LHIHIAEAPVAPVAKVGLADVTISGLKALQKGLHVEILPKYDCMQNDQVNNLKVLVD 481  
QY 61 MGEKVYFHSIKKGVHRVWIDHPFLAKVWGT--CSKLY--GPRSGADYL----- 107  
Db 482 VVQS--YF-----EGNLFNNKIWTGIVEGLPVTFIEPQHPAKFFWRAQYGEH 527  
QY 108 DNHRKRALCKAIEAARVLPFGGDCVFNVDHWSALVPLVKDEYQPKQGTAKSV 167  
Db 528 DFKRFAYSRALE--LLYQSKKIDIIHCHDQWTAFAVPLYWEAYANLG-FNSARIC 583  
QY 168 LAHNIAFGGRMWEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEE-----D 216  
Db 584 FTCHNPEYQG-----AQAQDLACCGLDVQQLD 611  
QY 217 EXPPLTGKTKYKINWLKGLIIAADKLVTSPNYATEIAADAGAGVGLDVTIRAKGTEGIV 276  
Db 612 REDMRDNDSHGRINNVKGAIVSNIVTIVPTALEVRSEGGKLODSLKHRSKTFVGL 671  
QY 277 NGMDIEWNPKTDKFLSAPYDQNSVYAGAAKAEALQAEIGL--PVDPAPLFAFIEGKLE 335  
Db 672 NGIDTDTWNPSTDRHLKVQYNDL--QGAANKAALRKLQNLSSSTNASQPLVGCITRLVP 730  
QY 336 QKGVDIILALPKLATPKVQIATIGTGAAYEKLVAIGTKYKRA-----KGVV 386  
Db 731 QKGVHLIRHAIYK-TAELGGQFVLLGSSVPFH-----IQREFEGADHFNQNNNRLRL 783  
QY 387 KESAPLAHMLTAGADFWLPSRPEPCGLIQLHAMHYGTVPVAVSTGGVLDTV----- 438  
Db 784 KYDSDLSHVIYASDMFTVPSMPEPCGLTQMIMRYGSPVIVKTKGGLNDSVDFDEDETI 843  
QY 439 -KEGVTGFHMGALNPDKLDEADADALAAVTRASEVFAAGRYPEMVAVNCISQDL-----SW 493  
Db 844 PKELRNGFTF--VHPDE-----KALSGAMERAFNY--NRKPEVNNKQLVQKMDRDFSW 893  
QY 494 SKPAQKWEGLLEEVV 508  
Db 894 ASSASQYEDIYQRAV 908

RESULT 7

```

US-10-163-214-12
; Sequence 12, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE NUMBER: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-163-214-12

```

Query Match	16.3%	Score 549;	DB 9;	Length 914;
Best Local Similarity	30.8%;	Pred. No. 1.1e-30;		
Matches 169;	Conservative 89;	Mismatches 202;	Indels 88;	Gaps 22;
QY	2	LDIVMVAEAVPWSKTGGGLDVTGGPIELVKRGHRVMTIAPRYD-QYADAWDTSVVVDI	60	
Db	421	LHIAHIAEAMAPVAKVGGGLADVTSGLGALKQKKGHLVEILPKYDCMQVDQVSNLKYLDV	480	
QY	61	MGEKVRYFHSIKGVHRVWIDHPWFLKAWGKTT--GSKLY--GPR-----SCADYLDNH	110	
Db	481	LVQS--YF-----EGMFMFNKIWTGTEGLPVYIEPQHPAMFFSRAQYIGE	526	
QY	111	---KPALECKAIEARVLPFGPGEDC--VFVANOWHSALVPVLLKDEYQPKQGFKA	166	
Db	527	DDFKRYSFSRAALE---LLYQSGKKVDIIHCHDWQTAFVAPLYWDVYANLG-FNSARI	581	
QY	167	VLAIHNIAFQGRMWEAEAFKDTKLPQAFDKLAFSDGVAKYVTEATPWEDEKPP-LTGKT	225	
Db	582	CFTCHNFYEQG-----TAPARDLAWC-----GLDVEHLDRDPRMRDNS	619	
QY	226	YKKINWLKGGIIAADKLVTVSPNIAFEIAADAGGVVELDTVIRAKGIEGIVNGMDEEWN	285	
Db	620	HGRINAVKGAUVYSNTVPTVPTALEVRSEGGRLQDTLKVHSRRFLGILNGIDTDTWN	679	
QY	286	PXTDFLSAPYDQNSVYAGKAAKAELQAEGL-PVDPTAPLEAFTGRLEEQKGVDIILIA	344	
Db	680	PCTDRYLKVQYNAKDL-QQKAANKAALREQNLNIAISAYSOPLVGCITRLVAQGVHLIRH	738	
QY	345	ALPKILATPKVQIATLGTGKAAYEKLNVNAGTKYKGRA-----KGWKFSAPLAHM	395	
Db	739	ALYK-TAELGGQFVLLGSSP-----VPIQREFEGIAHFQNNNNIRLLIKYDDALSHC	791	
QY	396	LTAGADEMLVPSPRFPCCGLILQHMHYGTVPVYASTGGVLVDY----KRGVTGFHMGALN	451	
Db	792	IYAASDMTIVPSIFEPCCGLQTMAMRYGSPVIRKTKGLNDSVDFDDDETIP---MEVRN	848	
QY	452	PDKLDEADADALAAVTRRASEVFAGGRYPEMWANCISQDL----SHSKFAQKWEGILLEV	507	
Db	849	GTTFVKADEQGLSSAMERAFNCYT--RKPEVWKQLVKQMDITDFSMDTSASQYEDIYQA	906	
QY	508	VYKGQVVA 515		
Db	907	VARARAVA 914		

RESULT 8  
US-10-163-214-2  
; Sequence 2, Application US/10163214  
; Publication No. US20030097688A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen M

```

: APPLICANT: Broglie, Karen E.
: APPLICANT: Butler, Catherine H.
: APPLICANT: Thorpe, Catherine J.
: TITLE OF INVENTION: Starch Synthase Isoform V
: FILE REFERENCE: BB1520 US NA
: CURRENT APPLICATION NUMBER: US/10/163,214
: CURRENT FILING DATE: 2002-06-05
: PRIOR APPLICATION NUMBER: 60/297,099
: PRIOR FILING DATE: 2001-06-08
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 2
: LENGTH: 909
: TYPE: PRT
: ORGANISM: zea mays
US-10-163-214-2

```

```

Query Match      16.2%  Score 545;  DB 9;  Length 909;
Best local Similarity 29.4%;  Pred. No. 2.1e-30;
Matches 159;  Conservative 92;  Mismatches 199;  Indels 90;  Gaps 20;

QY      2 LDIVMAAEVAPWSKGGGLGVTGGPIELVLRKGRHVMTIAPRYD-QYADAWTSTVVVDI 60
Db      416 LHVIVHAAEWAPYAKVGGGLADVLISGLKALOKKGLHVEILPKYICDMQHNOINNLKVLVD 475
QY      61 MGEKVRYFHSIKGVYHVRWIDHPWFIKAWWGKT--GSKLY--GPRSGADYL----- 107
Db      476 -----VVKSYFEGNNFANKIWGTGVEGLPVYFIEPQHPGKFFWRAQYYGEH 521
QY      108 DNHKREALFCKAAIEAARVLPEFGGEDC-VFVANDMHSALVPVLLKDEYQPKGQFTKAKS 166
Db      522 DDFKRSYFSRVALE---LLYOSGKKVDIIHCHDMQTAFVADLYWDVYANLG-FNSARI 576
QY      167 VLAHNIATQGRMWEBAFKDTKLPOAFAFDKLAFSDSGYKVVYATPEEDEKPPPLTKTY 226
Db      577 CFTCHNFXYQG-----IAPAQDLAYCGLDVVDHLDRPDR--MRDNSH 615
QY      227 KTNLWKGGLIADKLVTSPNYATEIAADAAGSGVELDITVIRAKGIEGLVNGMDIEWNP 286
Db      616 GRINVYKGAVVYINVTYSPYIAQVRSEGGGLQDTLKVHSKKFVGLIUNGIDTDTWNP 675
QY      287 KIDKFLSAPYDQNSVYAGKAAKEALQAEGLG-PVDPTAPLFAFIFGRLEEQKGVDIILAA 345
Db      676 STDRFLKVQYSANDLY-GKSANKAALRKQKLASTQASQPLVGCITRLVPQKGVHLIRHA 734
QY      346 LPKILATPKVQIALGLTGAAAY-EKLVNAIGTKYK--GRAKGVVYKFSAPLAHMLTAGADF 402
Db      735 IYKTELGG-QFVLLGSSPVQHLQRFEGFIADQFQNNNVRLKLDLDAHLAFIRASDM 793
QY      403 MLVPSFEPCGLIQLHAMHYGVVPVYASTGGLVDI-----VKEGVTGFHMGAL 450
Db      794 FIVPSFEPCGLTQVMARYGSPVYVRRITGGGLNDLVFDLDDETIPMEVRNGFTFL----- 848
QY      451 NPDKLDEADADALAAVVRRASEVFAGGRYPPEMVANCISODL-----SWSKPAQKWEGLEE 506
Db      849 ---KADEQD---FGNALERAFTNY--HRKPEVWKVQLVQDKMDIDFSNDTFSVSYEIIYQK 900

RESULT 9
US-10-163-214-13
; Sequence 13, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Broglie, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BB1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08

```



Db 100 YSENNVVFSAFAAAA--AAYLQADPAD--IVHLHDWHVCLLAGLLKNPLN-----VHS 151  
QY 165 KSVLAHINIAFOGHWTEAKDKLPQAAFDKLAFSGCIYAKVYTEATPMEDEKPLTGK 224  
Db 152 KIVFTIHNFGYRG-----YCSSTQLLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194  
QY 225 TYKKINWLGKGIITAADKLVTSVSPNYATEIAADAAGGVLELDTVIRAKG--IEGIVNGMDIE 282  
Db 195 L-----MKGALVCSDIITVSLTYVQEIINDKS--DYELHDAILARNSVFGIINGIDE 247  
QY 283 ENWPKTKFISAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFTGRLEEQ 336  
Db 248 VWNPKTDPALAVQVDSLLSEPDVLTFTKKEENRAVLVEKLGISSD-YFPLICVLSRIVEE 306  
QY 337 KG-----VDIILAAPKILATPKVQIATLGTGKAAYEKLVAIGKYGKGRAG-----VVK 387  
Db 307 KGPEFMKEIILHAMEHSYA-----FILIGTSQ--NEVLLNEFRNLQDCLASSPNIRLILD 359  
QY 388 FSAPLAHMLTAGADFMVPSRFPCCGLIQLHAMHYGTPVVPVASTGGGLVDTVKGVGTGFHM 447  
Db 360 FNDPLARTYAADMICIPSHREACGLTQLIAMRYGIVPLVRKTGLADIVPGVNGFTF 419  
QY 448 GALNPKDLDDADALAAATVRRASEVFAAGRYPEMVANCISODLSWSKPAQKWGLLEEV 507  
Db 420 --EDTNFNE-----FRAMLSNAV-----TYRQEPDVLNLLIESG 453  
QY 508 VYCKGGVATAKKEIKV 524  
Db 454 MLRASGLDAMAKHYVNL 470

## RESULT 12

US-10-163-214-10  
; Sequence 10, Application US/10163214  
; Publication No. US20030097686A1

## GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Thorpe, Catherine J.  
; TITLE OF INVENTION: Starch Synthase Isoform V  
; FILE REFERENCE: BB1520 US NA  
; CURRENT APPLICATION NUMBER: US/10/163,214  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/297,099  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-163-214-10

Query Match 11.4%; Score 385.5; DB 9; Length 293;  
Best Local Similarity 33.0%; Pred. No. 9.5e-20;  
Matches 101; Conservative 59; Mismatches 115; Indels 31; Gaps 9;  
QY 228 KINWLGKGIITADKLVTSVSPNYATEIAADAAGGVLELDTVIRAKGIEGIVNGMDIEWNP 287  
Db 1 RINAVKGAVVSYNIITVTSPTVALEVRSGRGLQDTLKVHSRKFGLGILNGIDTDWNP 60  
QY 288 TDKFISAPYDQNSVYAGKAAAEALQAEGL-PVDPATLFAFIEGRLEQKGVDTIILAL 346  
Db 61 TDRLYKVQYNAKDL--OGKAANKAAALREQLNLASYPQLVGCITRLVAQKGVHLIRHAI 119  
QY 347 PKILATPKVQIATLGTGKAAYEKLVAIGTKYKGRA-----KGVVKEFSAPLAHMLT 397  
Db 120 YK-TAELGQFVLLGSSP-----VPEIQREGEIADHFQNNNNIRLILKYDDALSHCIY 172  
QY 398 AGADFMVPSRFPCCGLIQLHAMHYGTPVVPVASTGGGLVDTV-----KEGVTGFHMGALNPD 453  
Db 173 AASUMFIVSFPEPCGLTQMIAMRYGSPVIRKVTGGLNDSVDFDEDETIP---MEVRNGF 229

QY 454 KLDEADALAAATVRRASEVFAAGRYPEMVANCISODL-----SWSKPAQKWGLLEEVY 509  
Db 230 TFVRADEGLSSAMERAFNCYT--RKPEVWKQLVKQMDTIDFSDTSASQIEDIYOKAVA 287  
QY 510 KGKGA 515  
Db 288 RARAVA 293

## RESULT 13

US-09-739-438-2

; Sequence 2, Application US/09739438  
; Patent No. US20020029394A1

## GENERAL INFORMATION:

; APPLICANT: Allen, Steve  
; APPLICANT: Beckles, Diane M.  
; APPLICANT: Thorpe, Cathy  
; TITLE OF INVENTION: Homologs of Starch Synthase DUI  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/739,438  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 60/171514  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-739-438-2

Query Match 6.4%; Score 215.5; DB 10; Length 117;  
Best Local Similarity 46.6%; Pred. No. 3.2e-08;

Matches 41; Conservative 17; Mismatches 25; Indels 5; Gaps 1;

QY 356 QIAILGTG-----KAAYEKLVAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRFE 410  
Db 8 QVVLGSA PDHRIQGDFTNLASKLHGEYHGRVKLCITYDEPLSHLIYAGADFILVPSMFE 67  
QY 411 PCGLIQLHAMHYGTPVVPVASTGGGLVDTV 438  
Db 68 PCGLTQLTAMRYGSIPIVRKTGLGYDTV 95

## RESULT 14

US-09-738-626-4738

; Sequence 4738, Application US/09738626  
; Publication No. US20020197605A1

## GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIALI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4738

138 GTEVDVLTVISQYTLRRFKSARFSGHPTFEHLPSGVVVKRFTPAT-----PED----- 184  
QY 305 KAAAKEALQAEGLPVDPTAPLFAFIFGRLEEQKGVDIILAAALPKILAT-PKQVIAILGTG 363  
Db 185 ----KSATRKKLGF--TDTPVIACNSRLVPRKQDLSLIKAMPQVIAARPDQLLIVGSG 238  
QY 364 KAAEKLVAIGTKYKRAKGVVVKFSAPLAHM----LTAGADFMVPSR-----FEPC 412  
Db 239 R--YESTLRLATDVSN-----VKFLGRLEYQDMINTLAAADIFAMPARTRGGLDVEGL 292  
QY 413 GLIQLHAMHYGTPVNVAST--GGLVDTVKESVTGFHMGALNPKDLDEADADALAATVRRAS 471  
Db 293 GIVYLEAQACG-VPVIAGTSGGAPETWTP-ATGLVVEGSDVDKLSSELLIELLDDPIRRAA 350  
QY 472 EVFAGGRYPMEVANCISQDLSWSKPAQKWEGJLE 505  
Db 351 MGAAGRAH-----VEAEWSWEIMGERLTNIQ 377

Search completed: June 4, 2003, 15:15:21  
Job time : 25.1047 secs

LENGTH: 409  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4738  
Query Match  
Best Local Similarity 27.2%; Score 206; DB 9; Length 409;  
Matches 83; Conservative 47; Mismatches 127; Indels 48; Gaps 15;  
QY 210 ATPMEDE--KPPLGKTKYKKNWL-KGIIAADKLIVSPNYATEIAADAAGGVELDTV 266  
Db 129 AHSLEPDRFWKEQLGGGVDVSWSEKNAMEYADAVIAVSARKDSILA-AYPRIEPDNV 187  
QY 267 IRAKGIEGVNGMDIEEWNPKTDKFLSAPYDQNSVYAGKAAAKEALQAEGLPVDPTAPL 326  
Db 188 -----RVVINGIDTELWQPR-----PFDD-----AEDSVLRSLG--VDPQRP 224  
QY 327 FAFIGRLEQKGVDIILAAALPKILATPKVQIAILGTGKAAYE--KLVNAIGTKYKRAKG 384  
Db 225 VAFVGRITRQKGVHELIKA--AALEFDESQVLVLCAGAPDTPETAARTTALVEELQAKREG 282  
QY 385 VVKFSAPLA-----HMLTAGADFMVPSRFEPCGLIQLHAMHYGTPVNVASTGGLVDTVK 439  
Db 283 IFWVQDMLGKDKIQEILTA-ADTFVCPSEIYEPGLIVNLEAMACNTAVASDVGGIPEVVV 341  
QY 440 EGVGTG--FHMALNPKDLDEADADA---LAATVRRASEVFAGGRYPMEVANCISQDLSWS 494  
Db 342 DGTGTALVHYDENDVETFERDIAEAVNKNVADRETAAKFGLAGR--ERAIN-----DFSNA 395  
QY 495 KPAQK 499  
Db 396 TIAQQ 400

RESULT 15  
US-09-738-626-5896  
Sequence 5896, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 5896  
LENGTH: 381  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5896  
Query Match  
Best Local Similarity 27.0%; Score 189; DB 9; Length 381;  
Matches 74; Conservative 43; Mismatches 95; Indels 62; Gaps 14;  
QY 260 GVELD-----TVIRAKG-----IEGVNGMDIEEWNPKTDKFLSAPYDQNSVYAG 304

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 08:28:25 ; Search time 3394.54 Seconds  
(without alignments)  
16743.883 Million cell updates/sec

Title: US-09-980-771A-4

Perfect score: 1953

Sequence: 1 gcgcgtgacatcgtgatggt.....ccgcaccocgcaagtcgcgc 1953

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.ov.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1953	100.0	1953	6	AX049325	Sequence
2	1953	100.0	2124	6	AX049323	Sequence
3	1953	100.0	3117	6	AX049322	Sequence
4	1951.4	99.9	3240	8	AF026420	Chlamydom
5	1953	81.6	1593	6	AX049329	Sequence
6	1314	67.3	1314	6	AX049327	Sequence
7	962.2	49.3	5856	8	AF433156	Chlamydom
8	536.8	27.5	2289	8	AF286320	Triticum
9	530	27.1	1812	8	AF486514	Hordeum v
10	530	27.1	2311	8	HVWAXYR	Hordeum v
11	528.4	27.1	1827	8	AF486515	Hordeum v
12	528.4	27.1	1827	8	AF486518	Hordeum v
13	525.6	26.9	2127	8	AF163319	Triticum
14	525.2	26.9	1812	8	AF486519	Hordeum v
15	525.2	26.9	1827	8	AF486516	Hordeum v
16	525.2	26.9	1827	8	AF486517	Hordeum v
17	524	26.8	1605	8	AF250137	Triticum
18	505.4	25.9	2028	8	AF113844	Triticum
19	505.2	25.9	2186	8	TAWAXYSS	Wheat waxy
20	505.2	25.9	2267	8	AY050174	Triticum
21	481.4	24.6	1801	8	AF113843	Triticum
22	475.8	24.4	2267	6	AK195560	Sequence
23	473.2	24.2	2115	8	AB089141	Setaria i
24	453.6	23.2	2287	8	SBU23945	Sorghum bic
25	449	23.0	1937	8	AB066093	Oryza sat
26	449	23.0	1937	8	AB066094	Oryza sat
27	447.4	22.9	2542	6	AB106491	Sequence
28	447.4	22.9	2542	8	GSWX	O.sativa wa
29	347	17.8	2168	8	MEGBSS	Sequence
30	340.2	17.4	2270	8	AF210699	Perilla f
31	329.6	16.9	2221	8	AB029546	Phaseolus
32	321.6	16.5	2211	8	AB071604	Ipomoea b
33	316.8	16.2	2161	6	AX394246	Sequence
34	313.2	16.0	2345	8	AMA6293	Antirrhin
35	312.8	16.0	3049	8	AF026421	Chlamydom
36	305.6	15.6	2124	8	IBU44126	Ipomoea bat
37	298	15.3	2092	8	AY094405	Arabidops
38	298	15.3	2115	8	AY088544	Arabidops
39	298	15.3	2192	8	AY123983	Arabidops
40	289.5	14.8	2107	8	AF097922	Astragalu
41	283.2	14.5	764	8	AF446083	Miscanthu
42	279.6	14.3	2035	8	PSSTARSYN	P.sativum m
43	274.2	14.0	2081	8	AF109395	Triticum
44	274	14.0	764	8	AF446082	Microsteg
45	273.2	14.0	721	8	AF446081	Cleistach

# ALIGNMENTS

RESULT 1  
LOCUS AX049325  
DEFINITION Sequence 4 from Patent WO0071734.  
ACCESSION AX049325  
VERSION AX049325.1 GI:12226092  
KEYWORDS synthetic construct.  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 1953)  
AUTHORS D'Hulst,C. and Ball,S.  
TITLE Starch granules containing a recombinant polypeptide of interest,  
method for obtaining same and uses  
JOURNAL Patent: WO 0071734-A 4 30-NOV-2000.

Pred. No. is the number of results predicted by chance to have a

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)  
Location/Qualifiers  
1. 1953  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="fragment de l'ADN complet codant la GBSSI de  
Chlamydomonas reinhardtii et codant la proteine GBSSI  
mature"  
CDS  
1. >1953  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC21781.1"  
/db\_xref="GI:12226093"  
/translation="ALDIIVMAAEVAPMSKGLGLDVTGGLPIELVKRHRVMTIAPR  
YDQYADNDSVVDIMGEKVPYFHSIKKGVHRVMDHPFLAKVKGKTSKLYGPRS  
GADYLDNRKALFCAAEARVLPFGEDCVFVANDMHSALVPVLLKDEYQPKGQ  
PLTKSALHAIHAFQGRWEEAFDKLPQAFADKLPFADGATKVPYVIRAKIEGIV  
PELTGKTKYNLWKGIIADKLIVSPNATEIAAAGGVDELDTVIRAKIEGIV  
NGMDTEENPKDKELSPADONSYAGKAKEALQAEGLPVDTPAFLEFIGLE  
EOKGVDIILALPKLILATPKVOIAILGTGKAAYEKLVNAIGTKYGRKGVYKFSAPL  
AHMLTAGADEFLVPSRFPCCGLIQLHMHYGVVNVASTGGLVDVKEGVTFHMGAL  
NDPKLDEADALAAITVRASEVFAAGRPYEMVANCISQDLSSWSPAKWEGILLEVV  
YKGGVATAKKEEIKVPVAEKIPGDLPAVSYAPNFIKPVSAVEGNGAAAPKVGITAP  
AMGAWRATTPSGPSAAATPKVTYKPAIPATAPKPTAGLKLAGEASTTSTSENGAAS  
NGNGNGASASKTSAKPLVSAATRKSA"  
BASE COUNT 344 a 712 c 608 g 289 t  
ORIGIN

Query Match 100.0%; Score 1953; DB 6; Length 1953;  
Best Local Similarity 100.0%; Pred. No. 1e-184;  
Matches 1953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCCTGGACATCTGATGGTTCGTCGAGTGCCTCCCTTGGTCCAGAGCGGGCGCTG 60  
DB 1 GCCTGGACATCTGATGGTTCGTCGAGTGCCTCCCTTGGTCCAGAGCGGGCGCTG 60  
QY 61 GCGGATGTGACTGGTGGCTGCTATTTAGCTGGTCAAGCGCGGCCACCGCGTCATGACC 120  
DB 61 GCGGATGTGACTGGTGGCTGCTATTTAGCTGGTCAAGCGCGGCCACCGCGTCATGACC 120  
QY 121 ATTTGCGGCTCGTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGGACATC 180  
DB 121 ATTTGCGGCTCGTACGACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGGACATC 180  
QY 181 ATGGCGAGAGGTCGCTACTTCCATCCATCAAGAGGCGGTGACCGCGTGTGATT 240  
DB 181 ATGGCGAGAGGTCGCTACTTCCATCCATCAAGAGGCGGTGACCGCGTGTGATT 240  
QY 241 GACCACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAGCTGTACGGCCCC 300  
DB 241 GACCACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAGCTGTACGGCCCC 300  
QY 301 CGCTCCGGCGCTGACTACCTGGACCAACCAAGCGCTTCGCCCTGTTCGAAAGCCGCT 360  
DB 301 CGCTCCGGCGCTGACTACCTGGACCAACCAAGCGCTTCGCCCTGTTCGAAAGCCGCT 360  
QY 361 ATTTGAGGCTGCGGCGTGTGCGCTTCGCGCCGCGGAGGACTCGCTTCCTGGTGGCCAAAC 420  
DB 361 ATTTGAGGCTGCGGCGTGTGCGCTTCGCGCCGCGGAGGACTCGCTTCCTGGTGGCCAAAC 420  
QY 421 GACTTGGCACTCCGCGCTTGGTGCCTGCTGCTGAAGGACGAGTACCGACCGCCAGGGCCAG 480  
DB 421 GACTTGGCACTCCGCGCTTGGTGCCTGCTGCTGAAGGACGAGTACCGACCGCCAGGGCCAG 480  
QY 481 TTCACCAAGCCAGTGGTGTGGCTATCCACAACTATCGCTTCCAGGGCCGATGTGG 540  
DB 481 TTCACCAAGCCAGTGGTGTGGCTATCCACAACTATCGCTTCCAGGGCCGATGTGG 540  
QY 541 GAGGAGGCTTTCAAGGACACGAGTGCCTCCAGGCGCGCTTTGACAAAGCTGGCTTCTCG 600  
DB 541 GAGGAGGCTTTCAAGGACACGAGTGCCTCCAGGCGCGCTTTGACAAAGCTGGCTTCTCG 600

QY 601 GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAGACCCCGG 660  
DB 601 GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAGACCCCGG 660  
QY 661 CTGACGGGAAAGACCTACAAGAAGATCACTGGCTGAAGGTTGGCATTTATCGCGCGGAC 720  
DB 661 CTGACGGGAAAGACCTACAAGAAGATCACTGGCTGAAGGTTGGCATTTATCGCGCGGAC 720  
QY 721 AAGCTGTGACTGTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATGCGCGCGGT 780  
DB 721 AAGCTGTGACTGTGTGCGCCCAACTACGCGACCGAGATCGCTGCCGATGCGCGCGGT 780  
QY 781 GTGGAGCTGGACACCGTATCTCCGCCCAAGGGCATTTAGGGCATTTGTAAACGCGATGGAC 840  
DB 781 GTGGAGCTGGACACCGTATCTCCGCCCAAGGGCATTTAGGGCATTTGTAAACGCGATGGAC 840  
QY 841 ATTGAGGAGTGGAAACCCCAAGACCGCAAGTTCCTGTCGCGCCTACGACCAAGACAGCAG 900  
DB 841 ATTGAGGAGTGGAAACCCCAAGACCGCAAGTTCCTGTCGCGCCTACGACCAAGACAGCAG 900  
QY 901 GTCTAGCGCGGCAAGGCGCGCCCAAGAGGCGCCTGACGCGCGAGCTGGGCGCTGCTGTG 960  
DB 901 GTCTAGCGCGGCAAGGCGCGCCCAAGAGGCGCCTGACGCGCGAGCTGGGCGCTGCTGTG 960  
QY 961 GACCCACACCGCGCGCTGCTGCTTCATCGCGCGCTGAGGAGGACAGAGGTTGTGGAC 1020  
DB 961 GACCCACACCGCGCGCTGCTGCTTCATCGCGCGCTGAGGAGGACAGAGGTTGTGGAC 1020  
QY 1021 ATCATCTGCGCGCTGCGCCCAAGATCTGCGCACCCCAAGGTGCAAGTGCCTCATCTG 1080  
DB 1021 ATCATCTGCGCGCTGCGCCCAAGATCTGCGCACCCCAAGGTGCAAGTGCCTCATCTG 1080  
QY 1081 GGTACCGGCAAGGCGCGCTACGAGAAGTGGTGAACGCCATCGGCACCAAGTCAAGGGC 1140  
DB 1081 GGTACCGGCAAGGCGCGCTACGAGAAGTGGTGAACGCCATCGGCACCAAGTCAAGGGC 1140  
QY 1141 CGCGCCAAAGGCGGTGTCAGTTCTCGCGCGCTTCCGCGCATGCTCACCGCGCGGCC 1200  
DB 1141 CGCGCCAAAGGCGGTGTCAGTTCTCGCGCGCTTCCGCGCATGCTCACCGCGCGGCC 1200  
QY 1201 GACTTCATGCTGTGTGCGCTTCCGCGCTGCGCATGCTCACCGCATGCTCACCGCGCATG 1260  
DB 1201 GACTTCATGCTGTGTGCGCTTCCGCGCTGCGCATGCTCACCGCATGCTCACCGCGCATG 1260  
QY 1261 CACTACGCTACCGTTCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCACCGCATG 1320  
DB 1261 CACTACGCTACCGTTCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCACCGCATG 1320  
QY 1321 GCGGTACCGGCTTCCACATGCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCAC 1380  
DB 1321 GCGGTACCGGCTTCCACATGCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCAC 1380  
QY 1381 GACGCGCTTGGCGCGCACCGTTCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCAC 1440  
DB 1381 GACGCGCTTGGCGCGCACCGTTCGCGCTTCCGCGCTTCCGCGCATGCTCACCGCATGCTCAC 1440  
QY 1441 GAGATGTGGCCAACTGCTATCAGCAGGACCTGCTTCTGGTCCAAAGCGCGCGCGCGCGG 1500  
DB 1441 GAGATGTGGCCAACTGCTATCAGCAGGACCTGCTTCTGGTCCAAAGCGCGCGCGCGCGG 1500  
QY 1501 GAGGCGCTGCTGGAGGAGTGTGTACGCGCAAGGCGCGCTGGCCACCGCCCAAGAGGAG 1560  
DB 1501 GAGGCGCTGCTGGAGGAGTGTGTACGCGCAAGGCGCGCTGGCCACCGCCCAAGAGGAG 1560  
QY 1561 GAGATCAAGTGTGCGGTTGCCGAGAGATCCCGCGCGACCTGCGCGCGCTGCTTCTTACGCC 1620  
DB 1561 GAGATCAAGTGTGCGGTTGCCGAGAGATCCCGCGCGACCTGCGCGCGCTGCTTCTTACGCC 1620  
QY 1621 CCCAACACCTTGAGACCGCTGCTCGGCTTCCGTTGGAGGCAACGCGCGCGCGCGCGCGCG 1680  
DB 1621 CCCAACACCTTGAGACCGCTGCTCGGCTTCCGTTGGAGGCAACGCGCGCGCGCGCGCG 1680  
QY 1681 GTGCGCACCG 1740



```
|||||
Db 1681 GTGCGACACCGCCCGCCATGGCGCTGGCGCGGAGACACCCCTCGGGCCCTCG 1740
QY 1741 CCGCGCGCGCCACCCGCAAGGTGACACCTACAAAGCCGCGCTGCGCGCACCGCCCAAG 1800
Db 1741 CCGCGCGCGCGACCCGCAAGGTGACACCTACAAAGCCGCGCTGCGCGCACCGCCCAAG 1800
QY 1801 CCCAAGACCGCTGGCTCAAGCTGGCGGTGAGGCTCCACACCTGACCTCGAGAAC 1860
Db 1801 CCCAAGACCGCTGGCTCAAGCTGGCGGTGAGGCTCCACACCTGACCTCGAGAAC 1860
QY 1861 GCGCTGCTCCAAAGCGCAAGCGGTGCTCGCGCTCCAAAGACCTCGGCTGCCAAG 1920
Db 1861 GCGCTGCTCCAAAGCGCAAGCGGTGCTCGCGCTCCAAAGACCTCGGCTGCCAAG 1920
QY 1921 CCCCTGCTCCCGCGCCACCGCAAGTCGCGC 1953
Db 1921 CCCCTGCTCCCGCGCCACCGCAAGTCGCGC 1953

RESULT 2
AX049323
LOCUS AX049323
DEFINITION Sequence 2 from Patent WO0071734.
ACCESSION AX049323
VERSION AX049323.1 GI:12226090
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.
REFERENCE
AUTHORS D' Hulst, C. and Ball, S.
TITLE Starch granules containing a recombinant polypeptide of interest,
method for obtaining same and uses
Patent: WO 0071734-A 2 30-NOV-2000;
JOURNAL CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)
FEATURES
source
1. .2124
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
note="fragment de la sequence compl te de l'DNAC codant
la GHSSI de Chlamydomonas reinhardtii."
1. .>2124
/note="unnamed protein product"
/codon_start=1
/translation="MAVASTSRPSSARPIVINAAGVYKKTANOLLRELARGSKST
SRSAVTGATGATCALDIVMAAEVAPWSTGGLGDTGGLPIELVYKRGHVRMTIAPRY
DOYADAWTSSVVDIMGEKRVYFHSIKKGVRHWIDHPWFLAKVWGTGSKLYGPRSG
ADYLDNHKREALFCKAAEAARVLFPFGEDCVFVANDWHSALVPLVLLKDEYOPKGF
TKASVLAHNTAFGRWEEAFKDTLPQAAFLDSDGYAKVYTEATPMEDEKP
PLTGKTIKWLGGIIIAADKLVTYSNPYATFIAADAAGVELDVTIRAKTEGIVN
GMIDIEWPKTDKLSAPYDQNSVYAGRAAAKEALQAEGLPVDPTAPLFAFIGLEE
QKGVDIILAAALPKIILATPKVOIILGTGKAAVEKLVNAGTGYKRAKGVYKESAPLA
HMLTAGADFMVLPSPREPCGLIQLHAMHVGTPVVASVGGIIVDTKVGTFHMGALN
PDKLDEADALAAATVRASEVFAGRTPPEWANCISQDLSWSKFAQKWEGLLEEVY
KGGVATRAKKEIKIPVAEKIPGDLPAVSYAPNTLKPVSASVEGGAAPKVGTTAPA
MGAWRATTPSGPSAAATPKVYTYKPAATKPAATKPAATKPAATKPAATKPAATKPAATK
GNNGASASKTSAAPKPLVSAATRKA"
BASE COUNT 369 a 773 c 656 g 326 t
ORIGIN

Query Match 100.0%; Score 1953; DB 6; Length 2124;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 1953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTGGACATGCTGATGTTGCTGCTGAGCTGCGCCCTTGGTCCAAAGCGGCGCCTG 60
Db 172 GCGCTGGACATGCTGATGTTGCTGCTGAGCTGCGCCCTTGGTCCAAAGCGGCGCCTG 231
QY 61 GCGCATGTGCTGGCTGCTGCTATTTAGCTGGTCAAGCGCGGCGCACCGCGCTGATGCC 120
|||||
```

```
|||||
Db 232 GCGCATGTGCTGGCTGCTGCTATTTAGCTGGTCAAGCGCGGCGCACCGCGCTCATGACC 291
QY 121 ATTGCCCTCGCTACGACCACTACCTGACGCTGGGACACCTCGGTGCTGTGACATC 180
Db 292 ATTGCCCTCGCTACGACCACTACCTGACGCTGGGACACCTCGGTGCTGTGACATC 351
QY 181 ATGGCGGAGAAAGTCCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGCTGTGGATT 240
Db 352 ATGGCGGAGAAAGTCCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGCTGTGGATT 411
QY 241 GACCAACCCCTGTTCTTGGCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCC 300
Db 412 GACCAACCCCTGTTCTTGGCAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGCGCCC 471
QY 301 CGCTCCGCGCTGACTACTTCCGACCAACCAAGCCCTTCGCGCCCTTCTTCAAGGCCGCT 360
Db 472 CGCTCCGCGCTGACTACTTCCGACCAACCAAGCCCTTCGCGCCCTTCTTCAAGGCCGCT 531
QY 361 ATTGAGGCTGCCCGCTGCTGCCCTTCGCGCCCGCGGAGGACTGCGTCTTCTGTGGCCCAAC 420
Db 532 ATTGAGGCTGCCCGCTGCTGCCCTTCGCGCCCGCGGAGGACTGCGTCTTCTGTGGCCCAAC 591
QY 421 CACTGGCACTCCGCGCTGGTGGCTGCTGCTGAAGGACGAGTACCAAGCCCAAGGGCCAG 480
Db 592 GACTGGCACTCCGCGCTGGTGGCTGCTGCTGAAGGACGAGTACCAAGCCCAAGGGCCAG 651
QY 481 TTCACCAAGGCAAGTGGTGTGCTATCCACACATCGGCTTCCAGGGCGCATGTGG 540
Db 652 TTCACCAAGGCAAGTGGTGTGCTATCCACACATCGGCTTCCAGGGCGCATGTGG 711
QY 541 GAGGAGGCTTTCAAGGACACCAAGCTGCCCGAGCGCGCTTTGACAAAGCTGCGCTTCG 600
Db 712 GAGGAGGCTTTCAAGGACACCAAGCTGCCCGAGCGCGCTTTGACAAAGCTGCGCTTCG 771
QY 601 GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCGCATGGAGGAGGACGAGAAGCCCGCG 660
Db 772 GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCGCATGGAGGAGGACGAGAAGCCCGCG 831
QY 661 CTGACGGGAAGACCTACAAGATCAACTGGCTGAAGGGTGCAATATCGCGCGCGAC 720
Db 832 CTGACGGGAAGACCTACAAGATCAACTGGCTGAAGGGTGCAATATCGCGCGCGAC 891
QY 721 AAGCTGCTGCTGTGCTGCCCAACTACGCGACCCGAGATCGCTGCCGATCGCGCGCGGT 780
Db 892 AAGCTGCTGCTGTGCTGCCCAACTACGCGACCCGAGATCGCTGCCGATCGCGCGCGGT 951
QY 781 GTGAGCTGACACCGCTATCCGCGCCCAAGGGCAATGAGGGCAATGTAAGCGCATGGAC 840
Db 952 GTGAGCTGACACCGCTATCCGCGCCCAAGGGCAATGAGGGCAATGTAAGCGCATGGAC 1011
QY 841 ATTGAGGAGTGGAAACCCCAAGACCGACAAGTCTCTGCTGCGCCCTACGACCAAGCAGC 900
Db 1012 ATTGAGGAGTGGAAACCCCAAGACCGACAAGTCTCTGCTGCGCCCTACGACCAAGCAGC 1071
QY 901 GTCTACCGCGCAAGGCGCGCCCAAGAGGCGCTTCGAGGCGCGAGCTGGGCGCTGCTGTG 960
Db 1072 GTCTACCGCGCAAGGCGCGCCCAAGAGGCGCTTCGAGGCGCGAGCTGGGCGCTGCTGTG 1131
QY 961 GACCCACCGCCCGCTGTTGCGCTTCATCGCGCGCGCTGGAGGACGAGAAGGGTGTGGAC 1020
Db 1132 GACCCACCGCCCGCTGTTGCGCTTCATCGCGCGCGCTGGAGGACGAGAAGGGTGTGGAC 1191
QY 1021 ATCATCTGGCGCGCTGCGCCCAAGATCCTGGCCACCCCAAGGTGCAGATCGCCATCTTG 1080
Db 1192 ATCATCTGGCGCGCTGCGCCCAAGATCCTGGCCACCCCAAGGTGCAGATCGCCATCTTG 1251
QY 1081 GSTACCGCAAGGCGCGCTTACGAGAAGCTGGTGAACCGCATCGGCACCAAGTACAAGGGC 1140
Db 1252 GSTACCGCAAGGCGCGCTTACGAGAAGCTGGTGAACCGCATCGGCACCAAGTACAAGGGC 1311
QY 1141 CGGCGCAAGGCGGTGTGAAGTTCTCGCGCGCGCTGGCGCACATGCTACCGCGCGCGC 1200
|||||
```

Db 1312 CGCGCAAGGGCGTGGTCAAGTTCTCGCGCGCCCTGGGCGACATGCTCACCGCCGCGCC 1371  
QY 1201 GACTTCATGCTGGTGGCTCGGCTTCGAGCCCTCGGCTCATCAGCTGACGCCATG 1260  
Db 1372 GACTTCATGCTGGTGGCTTCGAGCCCTCGGCTGATCAGCTGACGCCATG 1431  
QY 1261 CACTACGGTACCGTGGCTGGTGTAGCTCTCCACGGCGGCTGGTGGACACCGTCAAGGAG 1320  
Db 1432 CACTACGGTACCGTGGCTGGTGTAGCTCTCCACGGCGGCTGGTGGACACCGTCAAGGAG 1491  
QY 1321 GCGCTACCGGCTTCACATGAGGCGGCTGAAACCCGAGCAAGCTGACGAGCTGACGCC 1380  
Db 1492 GCGCTACCGGCTTCACATGAGGCGGCTGAAACCCGAGCAAGCTGACGAGCTGACGCC 1551  
QY 1381 GACGCCCTGGCGCCACCGTGGCGGCTGACGAGGCTTTGCGGGCGGCGCTACCC 1440  
Db 1552 GACGCCCTGGCGCCACCGTGGCGGCTGACGAGGCTTTGCGGGCGGCGCTACCC 1611  
QY 1441 GAGATGGTGGCCAACTGCATCAGCCAGGACCTGTCTGTGTCCTCAAGCCGCCAGAGTGG 1500  
Db 1612 GAGATGGTGGCCAACTGCATCAGCCAGGACCTGTCTGTGTCCTCAAGCCGCCAGAGTGG 1671  
QY 1501 GAGGGCTCTGGAGAGGTGTGTACGCAAGGGCGGCTGGCGACCGCCAGAGGAG 1560  
Db 1672 GAGGGCTCTGGAGAGGTGTGTACGCAAGGGCGGCTGGCGACCGCCAGAGGAG 1731  
QY 1561 GAGATCAAGTGGCCGTTGCGAGAGATCCCGGGGACCTGCGCGCGTGTCTACGCC 1620  
Db 1732 GAGATCAAGTGGCCGTTGCGAGAGATCCCGGGGACCTGCGCGCGTGTCTACGCC 1791  
QY 1621 CCCAACACCTGAAGCCGCTGTCCGCTCGGTGAGGGCAAGCGCGCGGCCCAAG 1680  
Db 1792 CCCAACACCTGAAGCCGCTGTCCGCTCGGTGAGGGCAAGCGCGCGGCCCAAG 1851  
QY 1681 GTGGGACACCGCCCGCATGCGGCTGCGGCGGCGGACACCGCCCTCGGCGCCCTCG 1740  
Db 1852 GTGGGACACCGCCCGCATGCGGCTGCGGCGGCGGACACCGCCCTCGGCGCCCTCG 1911  
QY 1741 CCGCGCGCGCCACCGCCCAAGGTGACCTACCAAGCGCGCCCTGCGCGCGCCCAAG 1800  
Db 1912 CCGCGCGCGCCACCGCCCAAGGTGACCTACCAAGCGCGCCCTGCGCGCGCCCAAG 1971  
QY 1801 CCCAAGACCGCTGGCTCAAGCTGCGGCTGAGGCTTCCACACCTGCGACCTCGGAGAC 1860  
Db 1972 CCCAAGACCGCTGGCTCAAGCTGCGGCTGAGGCTTCCACACCTGCGACCTCGGAGAC 2031  
QY 1861 GCGCTGCTCCAAAGCAAGCGCAAGGCTGCGGCTTCCAAAGCTGCGGCTGCGCAAG 1920  
Db 2032 GCGCTGCTCCAAAGCAAGCGCAAGGCTGCGGCTTCCAAAGCTGCGGCTGCGCAAG 2091

RESULT 3  
AX049322  
LOCUS AX049322 3117 bp DNA linear PAT 12-JAN-2001  
DEFINITION Sequence 1 from Patent WO0071734.  
ACCESSION AX049322  
VERSION AX049322.1 GI:12226089  
KEYWORDS  
SOURCE Chlamydomonas reinhardtii.  
ORGANISM Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadales; Chlamydomonas.  
REFERENCE 1. (bases 1 to 3117)  
D'Huist, C. and Ball, S.  
Starch granules containing a recombinant polypeptide of interest,  
method for obtaining same and uses  
Patent: WO 0071734-A 1 30-NOV-2000;  
JOURNAL CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)  
FEATURES Location/Qualifiers

source 1..3117  
/organism="Chlamydomonas reinhardtii"  
/db\_xref="taxon:3055"  
BASE COUNT 599 a 994 c 1011 g 513 t  
ORIGIN  
Query Match 100.0%; Score 1953; DB 6; Length 3117;  
Best Local Similarity 100.0%; Pred. No. 9; le-185;  
Matches 1953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGCTGACATCGTGTAGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGAGCGGCGCGCTG 60  
Db 186 GCGCTGACATCGTGTAGTGTGCTGCTGAGGTGCGCCCTTGGTCCAAAGAGCGGCGCGCTG 245  
QY 61 GCGGTGTGACTGTGCTGCTGCGCTTATTGAGCTGGTCAAGCGGCGCCACCGCTCATGACC 120  
Db 246 GCGGTGTGACTGTGCTGCTGCGCTTATTGAGCTGGTCAAGCGGCGCCACCGCTCATGACC 305  
QY 121 ATTGCCCCCTGCTTACGACCAAGTACGCTGAGCCCTGGACACCTGCGTGGTGGACATC 180  
Db 306 ATTGCCCCCTGCTTACGACCAAGTACGCTGAGCCCTGGACACCTGCGTGGTGGACATC 365  
QY 181 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGATT 240  
Db 366 ATGGGCGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGATT 425  
QY 241 GACCAACCCCTGGTTCCTTGGCCAAAGTGTGGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 300  
Db 426 GACCAACCCCTGGTTCCTTGGCCAAAGTGTGGGCAAGACCGGCTCCAAAGCTGTACGGCCCC 485  
QY 301 CGCTCGGCGCTGACTTACCTGGACACCAAGCGCTTCGCCCTTGCCTCTGTCGAAGCGCGCT 360  
Db 486 CGCTCGGCGCTGACTTACCTGGACACCAAGCGCTTCGCCCTTGCCTCTGTCGAAGCGCGCT 545  
QY 361 ATTGAGGCTCGCGCGTGTGCTTCCGCGCGCGGCGAGGACTGCGCTTTCGTTGGGCCAAC 420  
Db 546 ATTGAGGCTCGCGCGTGTGCTTCCGCGCGCGGCGAGGACTGCGCTTTCGTTGGGCCAAC 605  
QY 421 GACTGGCACTCGCGCTGTGCGCGCTCCTGCTGAAGGACGAGTACCAAGCGCGGCGAG 480  
Db 606 GACTGGCACTCGCGCTGTGCGCGCTCCTGCTGAAGGACGAGTACCAAGCGCGGCGAG 665  
QY 481 TTACCAAGGCGCAAGTGTGCTGCTATCCACAACTATCGCTTCCAGGGCGCGATGCG 540  
Db 666 TTACCAAGGCGCAAGTGTGCTGCTATCCACAACTATCGCTTCCAGGGCGCGATGCG 725  
QY 541 GAGGAGCTTTCAAGGACACGAGCTGCGCGCGCGCTTTCGACAGCTGGCGCTTCTCG 600  
Db 726 GAGGAGCTTTCAAGGACACGAGCTGCGCGCGCGCTTTCGACAGCTGGCGCTTCTCG 785  
QY 601 GACGGCTATGCCAAGGTTTACACTGAGGCGCACCCCTATGGAGGAGGAGCAAGCGCCCG 660  
Db 786 GACGGCTATGCCAAGGTTTACACTGAGGCGCACCCCTATGGAGGAGGAGCAAGCGCCCG 845  
QY 661 CTGACGGGAAGACCTACAAGAATCAACTGCGCTGAAGGTTGGCATTTATCGCGCGGAC 720  
Db 846 CTGACGGGAAGACCTACAAGAATCAACTGCGTGAAGGTTGGCATTTATCGCGCGGAC 905  
QY 721 AAGCTGTGACTGTGCTGCGGCACTACGCGACCGAGATGCGTGGCGCGCGCGGT 780  
Db 906 AAGCTGTGACTGTGCTGCGGCACTACGCGACCGAGATGCGTGGCGCGCGCGGT 965  
QY 781 GTGAGCTGGACACCGCTATCCGCGCAAGGGCATTCAGGGCATTTGTAACGGCATGGAC 840  
Db 966 GTGAGCTGGACACCGCTATCCGCGCAAGGGCATTCAGGGCATTTGTAACGGCATGGAC 1025  
QY 841 ATTGAGGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCGCTACGACCAAGACG 900  
Db 1026 ATTGAGGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCGCTACGACCAAGACG 1085  
QY 901 GTCTACGCGCGCAAGGCGCGCGCAAGGCGCTGAGGCGGAGCTGGGCTGCTGTG 960  
Db 1086 GTCTACGCGCGCAAGGCGCGCGCAAGGCGCGCTGAGGCGGAGCTGGGCTGCTGTG 1145

QY	961	GACCCACCGCCCGCCCTGTTCCGCTTCATCGCGCCCTGGAGGAGCAGAGGTTGTCAC	1020
Db	1146	GACCCACCGCCCGCCCTGTTCCGCTTCATCGCGCCCTGGAGGAGCAGAGGTTGTCAC	1205
QY	1021	ATCATCTGCGCGCCCTGCCCCAAGATCCTGSCCAACCCCAAGTGCAGATGCCATCTG	1080
Db	1206	ATCATCTGCGCGCCCTGCCCCAAGATCCTGSCCAACCCCAAGTGCAGATGCCATCTG	1265
QY	1081	GGTACCGGCAAGCGCCCTAGGAGAAGCTGTGAACGCTATCGGCACCAAGTACAAGGGC	1140
Db	1266	GGTACCGGCAAGCGCCCTAGGAGAAGCTGTGAACGCTATCGGCACCAAGTACAAGGGC	1325
QY	1141	CGCGCCCAAGCGCTGTCAAGTCTCGCGCCCTGCGGCACATGCTCACCGCGCGGCC	1200
Db	1326	CGCGCCCAAGCGCTGTCAAGTCTCGCGCCCTGCGGCACATGCTCACCGCGCGGCC	1385
QY	1201	GACTTCATGCTGTGCTCGCGCTCGAGCCCTGCGGCTGATCCAGCTGCAGCCCATG	1260
Db	1386	GACTTCATGCTGTGCTCGCGCTCGAGCCCTGCGGCTGATCCAGCTGCAGCCCATG	1445
QY	1261	CACCTACGGTACCGCTGCGGCTGAGCTCCACCGCGCGCTGTCGACACCGTCAAGGAG	1320
Db	1446	CACCTACGGTACCGCTGCGGCTGAGCTCCACCGCGCGCTGTCGACACCGTCAAGGAG	1505
QY	1321	GGCGTCAACCGCTTCCACATGGCGCCCTGAACCCCGACAGCTGGACGAGGTGACGCC	1380
Db	1506	GGCGTCAACCGCTTCCACATGGCGCCCTGAACCCCGACAGCTGGACGAGGTGACGCC	1565
QY	1381	GACGCCCTGCGCCGACCGCTGCGCGCTGCGAGGAGTGTTCGCGCGCGCGTACCCC	1440
Db	1566	GACGCCCTGCGCCGACCGCTGCGCGCTGCGAGGAGTGTTCGCGCGCGCGTACCCC	1625
QY	1441	GAGATGGTGCCCAACTGCATCAGCAGGACCTGTCTGTGTCGAAGCCGCCCAAGTGG	1500
Db	1626	GAGATGGTGCCCAACTGCATCAGCAGGACCTGTCTGTGTCGAAGCCGCCCAAGTGG	1685
QY	1501	GAGGGCTGCTGAGGAGGTGTGTCGCGCAAGGCGCGCTGCGCCCGCCCAAGAGGAG	1560
Db	1686	GAGGGCTGCTGAGGAGGTGTGTCGCGCAAGGCGCGCTGCGCCCGCCCAAGAGGAG	1745
QY	1561	GAGATCAAGTGCCCGTTCGCGAAGAGATCCCGCGGACCTGCGCGCGCTGCTACGCC	1620
Db	1746	GAGATCAAGTGCCCGTTCGCGAAGAGATCCCGCGGACCTGCGCGCGCTGCTACGCC	1805
QY	1621	CCCAACACCTGAAGCCCGTGTGCGCTCGTGGAGGGCAACGCGCGCGCCGCCCAAG	1680
Db	1806	CCCAACACCTGAAGCCCGTGTGCGCTCGTGGAGGGCAACGCGCGCGCCGCCCAAG	1865
QY	1681	GTGCGCACCAACCGCCCGCCGATGGCGGTGGCGCGGACACCCCGCTCGGCGCCCTCG	1740
Db	1866	GTGCGCACCAACCGCCCGCCGATGGCGGTGGCGCGGACACCCCGCTCGGCGCCCTCG	1925
QY	1741	CCGCGCGCGCCACCGCCCAAGGTGACCTACAAGCCCGCCCTGCGCGCCACCGCCAAG	1800
Db	1926	CCGCGCGCGCCACCGCCCAAGGTGACCTACAAGCCCGCCCTGCGCGCCACCGCCAAG	1985
QY	1801	CCCAACACCGTGGCTCAAGCTGCGCGGTGAGGCGCTCCACACCTCGACCTCGGAGAAC	1860
Db	1986	CCCAACACCGTGGCTCAAGCTGCGCGGTGAGGCGCTCCACACCTCGACCTCGGAGAAC	2045
QY	1861	GGCGCTGCCTTCCAAACGGCAACGGCAACGGTGCCTCGCGCTTCCAAAGACCTCGGCGCAAG	1920
Db	2046	GGCGCTGCCTTCCAAACGGCAACGGCAACGGTGCCTCGCGCTTCCAAAGACCTCGGCGCAAG	2105
QY	1921	CCCTGGTCTCGCGCGCCACCGCAAGTCCGCC	1953
Db	2106	CCCTGGTCTCGCGCGCCACCGCAAGTCCGCC	2138
RESULT 4			
AF026420			
LOCUS			
		3240 bp	mRNA linear PLN 05-NOV-2001
		AF026420	
DEFINITION	Chlamydomonas reinhardtii granule-bound starch synthase I precursor (STA2) mRNA, complete cds.		
ACCESSION	AF026420		
VERSION	AF026420.3		
KEYWORDS	GI:16716334		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii		
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.		
AUTHORS	1 (bases 1 to 3240)		
	D'Hulst,C., Wattedbled,F., Ral,J.-P., Abel,G.J., Kossmann,J. and Ball,S.G.		
TITLE	Cloning of a cDNA encoding for the GBSI in the green alga Chlamydomonas reinhardtii		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3240)		
AUTHORS	D'Hulst,C., Abel,G.J., Kossmann,J. and Ball,S.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-SEP-1997) Genetique Microbienne, Lab. Chimie Biologique UMR11 CNRS, Univ. Sci. Tech. de Lille, Villeneuve d'Ascq Cedex 59655, France		
REFERENCE	3 (bases 1 to 3240)		
AUTHORS	D'Hulst,C., Wattedbled,F., Ral,J.-P., Abel,G.J., Kossmann,J. and Ball,S.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUN-2000) Groupe de Glycobiologie des Protistes, Laboratoire de Chimie Biologique, UMR 8576 du CNRS, Universite des Sciences et Technologies de Lille, Villeneuve d'Ascq Cedex 59655, France		
REMARK	Sequence update by submitter		
REFERENCE	4 (bases 1 to 3240)		
AUTHORS	D'Hulst,C., Wattedbled,F., Ral,J.-P., Abel,G.J., Kossmann,J. and Ball,S.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-NOV-2001) Groupe de Glycobiologie des Protistes, Laboratoire de Chimie Biologique, UMR 8576 du CNRS, Universite des Sciences et Technologies de Lille, Villeneuve d'Ascq Cedex 59655, France		
REMARK	Sequence update by submitter		
COMMENT	On Nov 5, 2001 this sequence version replaced gi:8573759.		
FEATURES	Location/Qualifiers		
source	1..3240		
	/organism="Chlamydomonas reinhardtii"		
	/strain="137C"		
	/db_xref="taxon:3055"		
	/clone="CD142"		
gene	1..3240		
	/gene="STA2"		
CDS	138..2264		
	/gene="STA2"		
	/EC_number="2.4.1.21"		
	/function="App-Glucose:alpha-1,4-D-glucan-4-alpha-D-glucosyltransferase"		
	/note="GBSI"		
	/codon_start=1		
	/product="granule-bound starch synthase I precursor"		
	/protein_id="AAC17969.3"		
	/db_xref="GI:16716335"		
	/translation="MAVASTSRPSSRPVINAAPFGVKKTANOLLRELARGSARKSTSRSAVTGATGATCALDIVMAAEVAPWSKTGGTGLDVTGGTGLPIELVKRGHVMITAPRYDOYADWDTSVVDIMGERVTFHSIKGVRHVWIDHPWFLAKVNGTKSKLYGRPSGADYLDNRKDFALFCAAEARVLPFGPGDCVFVANDWHSALVPVLLKDEYQPKQFTRAKSVLAIHNTAFGRMEEAFKDTKLPQAAPDLAFSDGAKVYVTEATPEEDEKPTKGTYYKIKNLKGGIIAADKLVTSPNATEIAADAAGGVELDVIIRAKIGITVINGMIDEEWPKTDKELSPVDONSIVAGKAAKEALQAEGLPVDPTAPLPAFIGLEEOKSVLILALPKILATPKYQIALLGTGKAAEKLVAIGTKYKRGKGVKVSAPLAHMLTAGDFMLVPSRFPGCLLIQHAMHYGTVPVASTGGTGLVDITKRGVTFGHMALNPKLDEADADALAAATVRASEVFGAGRPYEMVANCISQDLSWSKPAQKEWGLLEEVYKGVGATAKKEIKVPAEKIPCDLPVAPSNITLKPVSASVEGNAAPKVGTTAPAMGAWRATTPSGSPAAATPKVTTKYPATAPATKPTAGLKLAGEASTTSTSENGAASNGNGASASAKTSAAPLVSAATKSA"		
transit_peptide	138..308		
	/gene="STA2"		

mat_peptide		309..2261						
		/gene="STA2"						
		/product="granule-bound starch synthase I"						
BASE COUNT	625 a	1038 c	1044 g	533 t				
ORIGIN								
Query Match	99.98; Score 1951.4; DB 8; Length 3240;							
Best Local Similarity	99.98; Pred. No. 1.3e-184;							
Matches 1952; Conservative	0;	Mismatches	1;	Indels	0;	Gaps		
0;								
QY	1	GGCTGGACATGTCGATGTTGCTGCTGAGTTCGCCCTTGGTCCAAAGAGGGCGGCCCTG	60					
Db	309	GGCTGGACATGTCGATGTTGCTGCTGAGTTCGCCCTTGGTCCAAAGAGGGCGGCCCTG	368					
QY	61	GCGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACC	120					
Db	369	GCGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACC	428					
QY	121	ATTGCCCTCGCTACGACACGATGCTGACGCTGGGACACCTTCGCTGCTGGACATC	180					
Db	429	ATTGCCCTCGCTACGACACGATGCTGACGCTGGGACACCTTCGCTGCTGGACATC	488					
QY	181	ATGGGCGAAGTTCGCTACTTCCATCCATCAAGAAGGGCGTGCACCGCTGTGGATT	240					
Db	489	ATGGGCGAAGTTCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCTGTGGATT	548					
QY	241	GACCACTCGCTGCTGCGCAAGTCTGGGCAAGACCGGCTCCAAAGCTGTACGGCCCC	300					
Db	549	GACCACTCGCTGCTGCGCAAGTCTGGGCAAGACCGGCTCCAAAGCTGTACGGCCCC	608					
QY	301	CGCTCGGGCGTACTACTTGACAAACACAAAGCGCTTGGCCCTGTTCTGCAAGGGCGCT	360					
Db	609	CGCTCGGGCGTACTACTTGACAAACACAAAGCGCTTGGCCCTGTTCTGCAAGGGCGCT	668					
QY	361	ATTGAGGCTGCCCGCTGCTGCCCTTCGCGCCCGGCGAGACTGCGTCTTCGTGGCCAAAC	420					
Db	669	ATTGAGGCTGCCCGCTGCTGCCCTTCGCGCCCGGCGAGACTGCGTCTTCGTGGCCAAAC	728					
QY	421	GACTGGCACTCCGCCCTGGTGGCGCTGCTGCTGTAAGAGCAGTACCAAGGCCCAAGGGCCAG	480					
Db	729	GACTGGCACTCCGCCCTGGTGGCGCTGCTGCTGTAAGAGCAGTACCAAGGCCCAAGGGCCAG	788					
QY	481	TTCCACCAAGGCCAAGTTCGCTGCTGCTATCCAAACATCGCTTCCAGGGCCGCTATGG	540					
Db	789	TTCCACCAAGGCCAAGTTCGCTGCTGCTATCCAAACATCGCTTCCAGGGCCGCTATGG	848					
QY	541	GAGGAGGCTTTCAAGGACACGAGCTGCCAGGCGCTTTGACAAAGTGGCCTTCTCG	600					
Db	849	GAGGAGGCTTTCAAGGACACGAGCTGCCAGGCGCTTTGACAAAGTGGCCTTCTCG	908					
QY	601	GACGGCTATGCCAAGGTTTACACTGAGGCGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGG	660					
Db	909	GACGGCTATGCCAAGGTTTACACTGAGGCGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGG	968					
QY	661	CTGACGGGAAGACCTTACAGAAGATCACTGGCTGAAGGTTGGCATTTATCGCGCGGAC	720					
Db	969	CTGACGGGAAGACCTTACAGAAGATCACTGGCTGAAGGTTGGCATTTATCGCGCGGAC	1028					
QY	721	AAGCTGTGACTGTGCGCCCACTAGCGACGAGATCGCTGCCGATGCCGCGCGGT	780					
Db	1029	AAGCTGTGACTGTGCGCCCACTAGCGACGAGATCGCTGCCGATGCCGCGCGGT	1088					
QY	781	GTGAGCTGGACACCGCTATPCCGGCGCAAGGGCATTGAGGGCATTGTGAACGGCATGGAC	840					
Db	1089	GTGAGCTGGACACCGCTATPCCGGCGCAAGGGCATTGAGGGCATTGTGAACGGCATGGAC	1148					
QY	841	ATTGAGGAGTGAACCCCAAGACGCAATTCCTGCTGCGCCCTACGACCGAAGACAGC	900					
Db	1149	ATTGAGGAGTGAACCCCAAGACGCAATTCCTGCTGCGCCCTACGACCGAAGACAGC	1208					
QY	901	GTCTACCGCGCAAGSCCGGCCAAGAGGGCCCTGCAGCGCGAGCTGGGCCCTGCGCTGTG	960					
Db	1209	GTCTACCGCGCAAGSCCGGCCAAGAGGGCCCTGCAGCGCGAGCTGGGCCCTGCGCTGTG	1268					

RESULT 5

AX049329

LOCUS

AX049329

1593 bp

DNA

linear

PAT 12-JAN-2001

DEFINITION Sequence 8 from Patent WO0071734.  
ACCESSION AX049329  
VERSION AX049329.1 GI:12226096  
KEYWORDS synthetic construct.  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 1593).  
AUTHORS D'Hulst,C. and Ball,S.  
TITLE Starch granules containing a recombinant polypeptide of interest,  
method for obtaining same and uses  
JOURNAL Patent: WO 0071734-A 8 30-NOV-2000;  
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)  
FEATURES Location/Qualifiers  
source 1..1593  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="fragment de l'DNAC complet codant la GBSSI de  
Chlamydomonas reinhardtii"  
CDS <1..>1593  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC21783.1"  
/db\_xref="GI:12226097"  
/translation="ALDIYVMAEVAEVPWSKGGIDVTGGPIELVKKGRHVMIIAPR  
YDQADAMDTSVVDIMGEKRYFHSIKKGVHRVWIDHPFLAKVMKGTGSKLYGPRS  
GADYLDNHRFALECKAIEAARVLFPFGEDCVFVANDHSAIYVLLKDEYQPKQ  
PFLAKVIAINIAIQGRWMEAFKDKLPQAAEDKLFASDGYAKVYEAUPMDEDEK  
PPLKGTYYKIKNLWGLIIAADKLVTSPNVAIEADAAGGVLDTVIRAKLEGIV  
NGMDIEWNPRTKPLSAPYDQNSVYAKAAKEALQELGLPVDPTAPLFAFIRGLE  
EOKGVDIILALPKILATPKVQIAILGTGKAAEKLNAIGTKYGRAGKGVYKFSAPL  
AHMLTAGDMFVSRPEPCGLIQHAMHYITVPVASTGGVLVDIVKRGVTFPHMAL  
NPDKLDEADALAAATVRASEVFAAGGYPWVANCTSDLSWSKPAQKWGLLEEVV  
YKGGVATAKEELKVPVAEKIP"  
BASE COUNT 290 a 539 c 510 g 254 t  
ORIGIN  
Query Match 81.6%; Score 1593; DB 6; Length 1593;  
Best Local Similarity 100.0%; Pred. No. 4.3e-149;  
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGCTGGACATCGTGTCTGCTGAGGTGCGCCCTTGCTCCAAAGACGGGCGGCGCTG 60  
Db 1 GCGCTGGACATCGTGTCTGCTGAGGTGCGCCCTTGCTCCAAAGACGGGCGGCGCTG 60  
QY 61 GCGGATGTGACTGTGGGCTGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACC 120  
Db 61 GCGGATGTGACTGTGGGCTGCTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACC 120  
QY 121 ATTGCCCTCGCTACGACACGATGCTGACGCTGGGACACCTCGGTGCTGTGGACATC 180  
Db 121 ATTGCCCTCGCTACGACACGATGCTGACGCTGGGACACCTCGGTGCTGTGGACATC 180  
QY 181 ATGGCGGAGAGTCCGCTACTTCCACTCCATCAAGAGGGGCTGCACCGCTGTGGATT 240  
Db 181 ATGGCGGAGAGTCCGCTACTTCCACTCCATCAAGAGGGGCTGCACCGCTGTGGATT 240  
QY 241 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAAGTGTACGGCCCC 300  
Db 241 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAAAGTGTACGGCCCC 300  
QY 301 CGCTCCGGCGTGAATCTGACCAACACACAGCGCTTCGCCCTGTCTGTGAAGGCCGT 360  
Db 301 CGCTCCGGCGTGAATCTGACCAACACACAGCGCTTCGCCCTGTCTGTGAAGGCCGT 360  
QY 361 ATTAGGCTGCGCGCTGCTGCGCTTCGCGCCCGCGGAGGACTGGTCTCTGCTGGCCAC 420  
Db 361 ATTAGGCTGCGCGCTGCTGCGCTTCGCGCCCGCGGAGGACTGGTCTCTGCTGGCCAC 420  
QY 421 GACTGGCACTCGCGCTGCTGCGCTTCCTGCTGAAGGACGAGTACAGCCCAAGGCGCAG 480  
Db 421 GACTGGCACTCGCGCTGCTGCGCTTCCTGCTGAAGGACGAGTACAGCCCAAGGCGCAG 480

QY 481 TTCACCAAGGCCAAGTCGGTGGCTATCCACAAATCGCCTTCCAGGSCCGCATGTGG 540  
Db 481 TTCACCAAGGCCAAGTCGGTGGCTATCCACAAATCGCCTTCCAGGSCCGCATGTGG 540  
QY 541 GAGGAGGCTTTCAGGACACGAAAGTGC0000AGGCGCCCTTTCACAAGCTTGGCCTTCTCG 600  
Db 541 GAGGAGGCTTTCAGGACACGAAAGTGC0000AGGCGCCCTTTCACAAGCTTGGCCTTCTCG 600  
QY 601 GAGGCTATGCCAAGCTTTACACTGAGGGCCACCC0000AGGACGAGGACGAGAGCCCCCG 660  
Db 601 GAGGCTATGCCAAGCTTTACACTGAGGGCCACCC0000AGGACGAGGACGAGAGCCCCCG 660  
QY 661 CTGACGGAAAGACCTTACAGAGATCAACTGGCTGAAGGGTGGCATTATCGCGCGCGAC 720  
Db 661 CTGACGGAAAGACCTTACAGAGATCAACTGGCTGAAGGGTGGCATTATCGCGCGCGAC 720  
QY 721 AAGCTGGTACTGTGTGCGCCCAACTACGGGACCGAGATCGCTCCGATGCGCGCGCGGT 780  
Db 721 AAGCTGGTACTGTGTGCGCCCAACTACGGGACCGAGATCGCTCCGATGCGCGCGCGGT 780  
QY 781 GTGGAGCTGGACACCGCTCATCGCGCCAAAGGCAATTGAGGGCATTGTGAACGCATGGAC 840  
Db 781 GTGGAGCTGGACACCGCTCATCGCGCCAAAGGCAATTGAGGGCATTGTGAACGCATGGAC 840  
QY 841 ATTGAGGCTGGAACCCCAAGACCGGACAAAGTTCTCTGCTGCGCCCTACGACCAAGAACAGC 900  
Db 841 ATTGAGGCTGGAACCCCAAGACCGGACAAAGTTCTCTGCTGCGCCCTACGACCAAGAACAGC 900  
QY 901 GTCTAGCGCGGCAAGCGCCCGCGGCAAGGAGGCGCTTCGAGCGCGAGCTGGGCGCTGCTGTG 960  
Db 901 GTCTAGCGCGGCAAGCGCCCGCGGCAAGGAGGCGCTTCGAGCGCGAGCTGGGCGCTGCTGTG 960  
QY 961 GACCCACCGCGCCCGCTGTGCGCTTCATCGCGCGCTGAGGAGCAGAGGCTGTGGAC 1020  
Db 961 GACCCACCGCGCCCGCTGTGCGCTTCATCGCGCGCTGAGGAGCAGAGGCTGTGGAC 1020  
QY 1021 ATCATCTGGCGCCCTGCGCCCAAGATCTGCGCACACCCCAAGGTGCAGATCGCCATCTTG 1080  
Db 1021 ATCATCTGGCGCCCTGCGCCCAAGATCTGCGCACACCCCAAGGTGCAGATCGCCATCTTG 1080  
QY 1081 GGTACCGGCAAGCGCCCTTACGAGAAAGTGGTGAAGCCCATCGGCAAGTACAAGGCG 1140  
Db 1081 GGTACCGGCAAGCGCCCTTACGAGAAAGTGGTGAAGCCCATCGGCAAGTACAAGGCG 1140  
QY 1141 CCGCGCAAGGCGTGGTCAAGTTCGCGCGCCCTGCGGCACATGCTCACCGCGCGCGC 1200  
Db 1141 CCGCGCAAGGCGTGGTCAAGTTCGCGCGCCCTGCGGCACATGCTCACCGCGCGCGC 1200  
QY 1201 GACTTCATGCTGGTGGCTTCGAGCGCTTCGAGCGCTGCGGCGCTGATCCAGTGCACGCCATG 1260  
Db 1201 GACTTCATGCTGGTGGCTTCGAGCGCTTCGAGCGCTGCGGCGCTGATCCAGTGCACGCCATG 1260  
QY 1261 CACTAGGTTACGTCGCGCTGTAGCTCCACCGCGCGCTGGTGCACACCGTCAAGGAG 1320  
Db 1261 CACTAGGTTACGTCGCGCTGTAGCTCCACCGCGCGCTGGTGCACACCGTCAAGGAG 1320  
QY 1321 GCGCTCACCGGCTTCACATGGCGCGCTTGAACCCCGACAAAGCTGGACGAGGTGACGCC 1380  
Db 1321 GCGCTCACCGGCTTCACATGGCGCGCTTGAACCCCGACAAAGCTGGACGAGGTGACGCC 1380  
QY 1381 GAGCGCTGGCGCCACCGTGC0000AGGAGGCTGCTGCTGTCCAAAGCCCGCCAGAGTGG 1440  
Db 1381 GAGCGCTGGCGCCACCGTGC0000AGGAGGCTGCTGCTGTCCAAAGCCCGCCAGAGTGG 1440  
QY 1441 GAGATGGTGGCCAACTGCATACGAGGACCTGCTGCTGTCCAAAGCCCGCCAGAGTGG 1500  
Db 1441 GAGATGGTGGCCAACTGCATACGAGGACCTGCTGCTGTCCAAAGCCCGCCAGAGTGG 1500  
QY 1501 GAGGCGCTGCTGGAGGAGTGGTGTACGGCAAGGCGCGGTGGCCACCGCCCAAGAGGAG 1560  
Db 1501 GAGGCGCTGCTGGAGGAGTGGTGTACGGCAAGGCGCGGTGGCCACCGCCCAAGAGGAG 1560

QY 1561 GAGATCAAGGTCCCGTTGGTCCGAGAGATCCCC 1593  
 |||||||  
 Db 1561 GAGATCAAGGTCCCGTTGGTCCGAGAGATCCCC 1593  
 |||||||

RESULT 6  
 LOCUS AX049327 1314 bp DNA linear PAT 12-JAN-2001  
 DEFINITION Sequence 6 from Patent WO071734.  
 ACCESSION AX049327  
 VERSION AX049327.1 GI:12226094  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM  
 REFERENCE  
 AUTHORS 1 (bases 1 to 1314).  
 D'Hulst, C. and Ball, S.  
 TITLE Starch granules containing a recombinant polypeptide of interest,  
 method for obtaining same and uses  
 JOURNAL Patent: WO 0071734-A 6 30-NOV-2000;  
 CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)  
 FEATURES  
 Location/Qualifiers  
 1..1314  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="fragment de l'DNAC complet codant la GBSSI de  
 Chlamydomonas reinhardtii"  
 <1..>1314  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC21782.1"  
 /db\_xref="GI:12226095"  
 /translation="ALDIVMAAEVAPMSKGGDVTGGPILVLRHVRMTIAPR  
 YDQYADAWDTSVVDIMGEKRYFHSIKKGVHRYWIDHPWELAKVWGKTSKLYGPRS  
 GADYDNRKALFCKAEARVLPFGEDCVFVANDWHSALVFLIKETDEIQKQ  
 FTKAKSLAINTAFQGRWEEAFDKLPQAFDFKLAIFSAGYAKVYKTEIPMEDEK  
 PPLTKYTKIKLNMKGIIADKLIVTSNPYATETADAGSGVELDVIYRAKIGIIV  
 NMDLEENWPKTDFLSAPYQNSVYAKAAKEALQALGLPVDPTAPLFAFIRGLE  
 EOKGVDIILALPKPLTKVQIAILGTGKAAYEKLVAIGTKYGRKGVVKEFAPL  
 AHMLTAGDFMIVPSRFEPCGLIOLHAMHYTVPVVASTGGLVDIV"  
 BASE COUNT 238 a 449 g 219 t  
 ORIGIN

Query Match 67.3%; Score 1314; DB 6; Length 1314;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-121;  
 Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTGGACATCGTATGTTGCTGCTGAGTGCAGCCCTTGTCCAAAGACGGCGCCCTG 60  
 |||||||  
 Db 1 GCCTGGACATCGTATGTTGCTGCTGAGTGCAGCCCTTGTCCAAAGACGGCGCCCTG 60  
 |||||||

QY 61 GCGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACC 120  
 |||||||  
 Db 61 GCGATGTGACTGGTGGCTGCTATTGAGTGGTCAAGCGGGCCACCGCGTCATGACC 120  
 |||||||

QY 121 ATTGCCCTCGCTACGACACGACGTGACGCTGAGGACACCTCGGTGGTGGGACATC 180  
 |||||||  
 Db 121 ATTGCCCTCGCTACGACACGACGTGACGCTGAGGACACCTCGGTGGTGGGACATC 180  
 |||||||

QY 181 ATGGGCGAGAAGTCCGCTTACCTCCACTCAACAAGGGGTGACCGCGGTGTGGATT 240  
 |||||||  
 Db 181 ATGGGCGAGAAGTCCGCTTACCTCCACTCAACAAGGGGTGACCGCGGTGTGGATT 240  
 |||||||

QY 241 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGCTGACGGCCCG 300  
 |||||||  
 Db 241 GACCAACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGCTGACGGCCCG 300  
 |||||||

QY 301 CGCTCGGGCGCTGACHTACCTGGACACACACAGCGGTTCCGCTGTTCGCAAGGCCGCT 360  
 |||||||  
 Db 301 CGCTCGGGCGCTGACHTACCTGGACACACACAGCGGTTCCGCTGTTCGCAAGGCCGCT 360  
 |||||||

QY 361 ATTGAGGCTCCCGCGTGTGCTGCCCTTCGCGCCCGGGGAGGACTGCGTCTCTCGTGGCCAAC 420  
 |||||||

Db 361 ATTGAGGCTCCCGCGTGTGCTCCCTTCGGCCCGCGGAGGACTGCGTCTCTGTTGGCCAAC 420  
 |||||||

QY 421 GACTGGCACTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 |||||||

Db 421 GACTGGCACTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 |||||||

QY 481 TTACCAAGGCGCAAGTGGTGGTGTATCCCAACATGCGCTTCCAGGGCCGCGCATGTGG 540  
 |||||||

Db 481 TTACCAAGGCGCAAGTGGTGGTGTATCCCAACATGCGCTTCCAGGGCCGCGCATGTGG 540  
 |||||||

QY 541 GAGGAGGCTTTCAAGGACAGAGTGGCCCGAGCGCGCTTTTGACAAGCTGGCTTCTCG 600  
 |||||||

Db 541 GAGGAGGCTTTCAAGGACAGAGTGGCCCGAGCGCGCTTTTGACAAGCTGGCTTCTCG 600  
 |||||||

QY 601 GAGGCTATGCCAAGGTTTACACTGAGGCGCACCCCGCATGAGGAGGACGAGAACGCCCGG 660  
 |||||||

Db 601 GAGGCTATGCCAAGGTTTACACTGAGGCGCACCCCGCATGAGGAGGACGAGAACGCCCGG 660  
 |||||||

QY 661 CTGACGGGAAAGACCTACAAGAGATCAACTGGCTGAAGGGTGGCATTTATCGCCGCGAC 720  
 |||||||

Db 661 CTGACGGGAAAGACCTACAAGAGATCAACTGGCTGAAGGGTGGCATTTATCGCCGCGAC 720  
 |||||||

QY 721 AAGCTGGTACTGTGTGCGCCAACTAGGCGACCGAGATCGCTCCGATGCCCGCGCGGT 780  
 |||||||

Db 721 AAGCTGGTACTGTGTGCGCCAACTAGGCGACCGAGATCGCTCCGATGCCCGCGCGGT 780  
 |||||||

QY 781 GTGGAGCTGGACACCGTGTATCCGCGCAAGGGCATTTGAGGGCATTTGTAACGCGCATGGAC 840  
 |||||||

Db 781 GTGGAGCTGGACACCGTGTATCCGCGCAAGGGCATTTGAGGGCATTTGTAACGCGCATGGAC 840  
 |||||||

QY 841 ATTGAGGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 900  
 |||||||

Db 841 ATTGAGGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGACAGC 900  
 |||||||

QY 901 GTCTACGCGCGCAAGCGCGCCGCCCAAGAGGGCCCTTCAGGCCCGAGCTGGCCCTGCTGTG 960  
 |||||||

Db 901 GTCTACGCGCGCAAGCGCGCCGCCCAAGAGGGCCCTTCAGGCCCGAGCTGGCCCTGCTGTG 960  
 |||||||

QY 961 GACCCACACCGCCCGCTGTTGCGCTTTCATCGCGCGCTGAGGAGGACAGAGGGTGTGGAC 1020  
 |||||||

Db 961 GACCCACACCGCCCGCTGTTGCGCTTTCATCGCGCGCTGAGGAGGACAGAGGGTGTGGAC 1020  
 |||||||

QY 1021 ATCATCTGCGCGCTGCGCCAAAGTTCCTGGCCACCGCCCAAGGTGCAGATCGCCATCCTG 1080  
 |||||||

Db 1021 ATCATCTGCGCGCTGCGCCAAAGTTCCTGGCCACCGCCCAAGGTGCAGATCGCCATCCTG 1080  
 |||||||

QY 1081 GGTACCGCAAGCGCGCTTCAAGTTCCTGGCGCCCTTCGCGCACATGCTCACCGCGCGCC 1140  
 |||||||

Db 1081 GGTACCGCAAGCGCGCTTCAAGTTCCTGGCGCCCTTCGCGCACATGCTCACCGCGCGCC 1140  
 |||||||

QY 1141 CGCGCCAAAGGGCGTGGTCAAGTTCCTGGCGCCCTTCGCGCACATGCTCACCGCGCGCC 1200  
 |||||||

Db 1141 CGCGCCAAAGGGCGTGGTCAAGTTCCTGGCGCCCTTCGCGCACATGCTCACCGCGCGCC 1200  
 |||||||

QY 1201 GACTTCATGCTGTGCTGCGCTTCGAGCCCTTCGCGCCCTTCGAGCTGCAGCGCATG 1260  
 |||||||

Db 1201 GACTTCATGCTGTGCTGCGCTTCGAGCCCTTCGAGCCCTTCGAGCTGCAGCGCATG 1260  
 |||||||

QY 1261 CACTACGCTACCGTGGCGCTGTGAGGCTCCACCGCGCGCTGCTGCACACCGTTC 1314  
 |||||||

Db 1261 CACTACGCTACCGTGGCGCTGTGAGGCTCCACCGCGCGCTGCTGCACACCGTTC 1314  
 |||||||

RESULT 7  
 LOCUS AF433156 5856 bp DNA linear PLN 23-AUG-2002  
 DEFINITION Chlamydomonas reinhardtii granule-bound starch synthase I (STA2)  
 gene, complete cds.  
 ACCESSION AF433156  
 VERSION AF433156.1 GI:16755882  
 KEYWORDS Chlamydomonas reinhardtii.











```
QY 929 AGGCCTCGAGCCGAGCTGGGCTGCTGTGTAACCCACCACCCGCTGTTGCGCTTCA 988
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1136 AGGCCTCGAGCCGAGCTGGGCTGCGGTGGACAGGAGTGGCGTGGCTTCA 1195
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 989 TCGGCGGCTCGAGGAGCAGAGGCTGGGACATCATCTGCGCGCCCTGCGCAAGATCC 1048
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1196 TCGGCGGCTCGAGGAGCAGAGGCTGGGACATCATCTGCGCGCCCTGCGCAAGATCC 1255
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1049 TGGCCACCCCAAGTGCAGATCGCATCTCTGGGTACCGGCAAGCGCGCTACGAGAAGC 1108
||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1256 TGAAGGAGGAGACCTCCAGATCATCTCTTGGCACCGGGAAGAGTTGAGAAGC 1315
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1109 TGGTGAACCCATCGCACCAAGCAAGGCGCGGCCCAAGGCGCTGGTCAAGTTCTCGG 1168
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1316 TGCTCAAGAGCATGAGGAGAGATTCCCGGCAAGTGAGGCGCGGTGAGTTCAAGC 1375
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1169 CGCCCTCGGCGACATGCTCACCGCGCGCGGCACTTCAATGCTGCTGCGCTTGG 1228
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1376 CGCGGCTAGCTACCAAGATGATGCGCGCGCGGCACTTGTGCTGTCACCAAGCGCTCG 1435
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1229 AGCCCTGCGGCTGATCCAGCTGCAGCGCATGCTACGCTACCGTGCCTGCTAGCCT 1288
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1436 AGCCCTGCGGCTGATCCAGCTGCAGCGCATGCTACGCTACCGTGCCTGCTAGCCT 1495
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1289 CCACCGCGGCTGCTGCGACACCGTCAAGGAGGCGCTCACCGGCTTCCACATGGCGGCC 1348
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1496 CCACCGCGGCTGCTGCGACACCGTCAAGGAGGCGCTCACCGGCTTCCACATGGCGGCC 1555
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1349 TGA-----ACCCGACAGCTGAGGAGCTGACCGCGAGCGCCCTGGCGGCGACGCTGC 1402
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1556 TCAGCGTCACTCAACGCTGGTGGAGCGCGGCGAGCTGAAGAGGTGGCGACCACTGA 1615
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1403 GCGTGCACGAGGCTTGGCGGCGCGCTACCGGAGATGGTGGCAACTGCATCA 1462
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1616 AGCGGCGCTCAAGCTGCTCGGACCGCGGCTACCAAGAGATGGTCAAGAACTGCATGA 1675
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1463 GCAGGACCTGCTCTGCTGCAACCGCGCGGAGAGTGGGAGGCGCTGCTGGAGAGGTGG 1522
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1676 TCCAGGATCTCTCTGGAAGGACCTGCCAAGAACTGGGAGGACGCTTCTTGGAACTGG 1735
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 1523 TGACGCAAGGCGGCTGGCCACCGCAAGAGAGAG 1560
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1736 GGTGGAGGAGCGAGCGCGGAGTGTGCGGAGGAG 1773
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 10
HWAXYR
LOCUS HWAXYR 2311 bp mRNA linear PLN 12-SEP-1993
DEFINITION Barley mRNA pcwx27 for waxy locus.
ACCESSION X07932
VERSION 1
KEYWORDS starch synthase; waxy locus.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticaceae; Hordeum.
REFERENCE 1 (bases 1 to 2311)
Zohde,W
AUTHORS Direct Submission
TITLE Submitted (21-JUN-1988) Zohde W., Max Planck Institute, Erwin Baur
JOURNAL Institut, D-5000 Koln 30, FRG
REFERENCE 2 (bases 1 to 2311)
Rohde,W., Becker,D. and Salamini,F.
AUTHORS Structural analysis of the waxy locus from Hordeum vulgare
TITLE Nucleic Acids Res. 16 (14B), 7185-7186 (1988)
JOURNAL 88303345
MEDLINE 2970062
PUBMED
COMMENT The sequence overlaps with the genomic sequence of the waxy locus,
x07931, from bp 1918 to bp 2301.
FEATURES
source 1..2311
```

```
/organism="Hordeum vulgare subsp. vulgare"
/strain="Vogelsanger Gold"
/db_xref="taxon:112509"
/map="chromosome 1"
/clone="pcwx27"
/clone_lib="EMBL4"
CDS
226..2037
note="starch synthase"
/codon_start=1
/protein_id="CAA30756.1"
/db_xref="GI:19129"
/db_xref="SWISS-PROT:P09842"
/translation="MAALATSLATSGTVLGVYDRFRPFGFGLRPNPADALGMRT
IGASAAKPSKRAHRSRCLSVVVSATSGSMNLVFGAEMAPWSTKGLGDRVLGSLP
PAMAANGHRVWVSPRYDOYKADWTSIETKVADEYERAFPHCYTKRGVDRVIDH
PWFLEKVRGKTKETIYGPDAQDYNQOYRSLQCAALEAPRIILNNPNYFSGPYG
EDVWPCNDWHTGLLACYLKSNTQYRTAKVAFCHNISYQGRFSDFDFAQLNLP
DRFKSDFDIDGDKPVEGRKINWKAGILOADKVLTVSPYAEELISEAARGCELDN
IMRLGITGIVNGMDVSDWPTKDFLAVNVDITTALEAKALKKELKALOEAGLVPVDRK
VPLVAFIGRLERQKGPDMIAAIPILKEEDVQIILLCTGKKFKELKALOEAGLVPVDRK
VRVVRFNAPLAHOMMAGADLLAVSRFEPCGLIQLOQWRYGTGTPCVCASTGSGIVTIV
EGKTFHMRGLSVDENVVPADVKVATTLKRAVKVGPAYQEMVKNCMIOQLSNWK
PAKNVEDVLELGVSESGPIVGEIPIAMENVAAP"
2276
note="poly A site (pcwx4)"
2283
note="poly A site (pcwx27)"
2301
note="poly A site (pcwx18)"
BASE COUNT 501 a 688 c 714 g 403 t
ORIGIN
Query Match 27.1%; Score 530; DB 8; Length 2311;
Best Local Similarity 61.6%; Pred. No. 4.7e-44;
Matches 985; Conservative 0; Mismatches 520; Indels 93; Gaps 5;
QY 5 TGCACATCTGATGCTGCTGAGTCCGCTTGGTCCCAAGACGCGGCTGGCG 64
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 452 TGRACCTCGTGTGTCGCGCGCGAGATGGCGCTCGAGACAGACCGCGCTTGGCG 511
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 65 ATGTGACTGTGGCTGCTGCTATTGAGCTGATCAAGCGCGGCGACCGGCTGATGACCATG 124
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 512 ATGTCTCTGCGCGGCTCCCGCGCCATGGCGGCAACGCTACCGGCTGATGCTGCTCT 571
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 125 CCGCTCGCTAGCACCATGCTGACCGCTGGGACACTCGGTGCTGTGACATCATGG 184
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 572 CCGCGCTACGATCAGTCAAGACGCGCTGGGACACGAGCTATCTCCGAGATCAAG 631
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 185 -----GGGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGCGCTGACCGCG 232
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 632 TCCTGACGAGTACGAGGCTGAGTTCCTCCACTGCTACAGCGGAGTGGACCGCG 691
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 233 TGTGATTTGACACCCCTGCTGCTGCGCAAGTCTGCGGCAAGACCGGCTCCAGCTGT 292
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 692 TGTTCATCGACACCCCTGCTGCTGGAAGAGTCCGCGGCAAGACCAAGAGAAGATCT 751
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 293 ACGGCCCCGCTCGGCGCTGCTACCTGGGACACCAAGCGCTGCGCTGCTTCTGCA 352
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 752 ACGGCCCCGACCGCGGACAGACTATGAGGACCAACGAGCGCTTACGCTTCTGCCC 811
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 353 AGGCGCTATTGAGGCTGCGCGCTGCTGCGCTTC-----G 388
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 812 AGGACGACTCGAGCACCCAGGATCTCAACCTCAACAACACCCCTTCTTCTGCTC 871
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 389 GCGCGCGAGGACTGGTCTTGTGGCCAAAGACTGGCACTCGCGCTGTCGCGCTCC 448
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 872 CTTACGGGAAGAGCTGGTGTCTGTGCAACGACTGGCACACGCGGCTTCTGCGCTCT 931
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 449 TGTGAGGACGAGTACAGCCCAAGGCGGCTTCAACCAAGGCGGAGTGTGCTGGTCA 508
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 932 ACCTCAAGAGCAACTACAGTCCCAATGGCATCTACAGGACGCGGAGGTGGCTTCTGCA 991
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 509 TCCACAACATGCGCTTCCAGGCGCGCATGTGGGACGAGGCTTTCAAGGACACAGAGCTGC 568
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```

|||||
992 TCACAACTCTCGTACAGGCGCGTTCCTTCGAGACTTTGGCCAGCTCAACCTGC 1051
QY CCAGGCGCGCTTTGACAGCTGGCTTCGAGAGCGCTATGCCAAGTTTACACTGAGG 628
Db |||||
1052 CCAGAGGTTCAAGTCGTCTTCGACTTATTCAGCGCTACGACAAG----- 1098
QY CCACCCCATGGAGGAGGACGAGAACGCCCGCGTGCAGCGGAAGACCTACAGAAGATCA 688
Db |||||
1099 -----CCGTGGAGGGCG-----AAGATCA 1120
QY ACTGGCTGAAGGTGGCAATTATCCCGCGACAAAGCTGTGACTGTCTGCCCAACTACG 748
Db |||||
1121 ACTGGATGAAGCGCGGATCCTCAGCGCGACAAAGTGTGACGGTGCAGCCCTACTACG 1180
QY CGACCCAGATCGCTCGCGATCCCGCGGGGTGTGAGCTGGACACCGTCATCCGCGCCA 808
Db |||||
1181 CTGAGGAGCTCATCTCTGGGAAGCGGAGGCTGCGAGCTCGACAACATCATCGGCTCA 1240
QY AGGCATTGAGGCGCATTTGAAGCGCATGCACATTGAGGAGTGAACCCCAAGCGACA 868
Db |||||
1241 CGSGGATCACCGGAATCGTCAAGCGCATGACCTCAGTGAAGGACCCCAAGGACA 1300
QY AGTTCCTGTCGCGCTCAACTACGACATCACCAACCGGTTGGAGGCCAAGCGCTGAACAG 928
Db |||||
1301 AGTTCCTGTCGCGCTCAACTACGACATCACCAACCGGTTGGAGGCCAAGCGCTGAACAG 1360
QY AGGCGCTGACGCGGAGTGGGCTGCTGACATCATCTGCGCGCGCTGCGCCCAAGTCC 1048
Db |||||
1421 TCGGCGAGGTGAGGAGCAGAGGCGCGGAGTGAATCGCGCGCATCCCGGAGATCC 1480
QY TGGCACCCCGCAAGTGCAGATCGCCATCTCTGGTACCAGGCGCGCTACGAGAAGC 1108
Db |||||
1481 TGAAGGAGGAGGACGCTCCAGATCATCTCTGGTACCAGGCGCGCTACGAGAAGC 1540
QY TGGTGAACGCCATCGCCACCAAGTACAGGCGCGCGCGCGCGGCGGTGCTCAAGTCTCGG 1168
Db |||||
1541 TGCTCAAGAGCATGAGGAGGAAGTTCCCGGGAAGGTGAGGCGCGCTGTCAGGTTCAAGC 1600
QY GCGCCCTGGCGCACATGCTCACCGCGCGCGCGCTTATCTGTTGCGCTCGCGCTGCG 1228
Db |||||
1601 GCGCGCTAGCTCACAGATGATGCGCGCGCGCGCTTCTGCTGCTGCTCACCGCGCTCG 1660
QY AGCCCTGCGGCTGATCCAGCTGACGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1288
Db |||||
1661 AGCCCTGCGGCTGATCCAGCTGACGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720
QY CCACGCGCGCTGCTGACACCGCTCAAGGAGGCGCTCACCGGCTTCCAGTGGCGGCC 1348
Db |||||
1721 CCACGCGCGGCTGCTGACACCGCTGCTGAGGAGGCAAGACCGGCTTCCAGTGGCGGCC 1780
QY TGA-----ACCCGACAAAGTGGACGAGGCTGACGCGCGCGCGCTGCGCGCGCGCTGCG 1402
Db |||||
1781 TCAGCGTGCAGTGCACGCTGTGGAGCGCGGAGCTGAAGAGTGTGGGACCCCTGA 1840
QY GCGTGCAGGAGGTGTTGCGGCGCGCGCTACCCGAGATGTTGGCCCAACTCATCA 1462
Db |||||
1841 AGCGGCGCGTCAAGTGTCTGCGCGCGCGCGCTACAGGAGATGTTCAAGAACTCATCA 1900
QY GCGAGGACTGCTCTGTTGTTGCGGCGCGCGCTACCCGAGATGTTGGCCCAACTCATCA 1522
Db |||||
1901 TCCAGGATCTCTCTGGAAGGACCTGCAAGAACTGGGAGGACGTGCTTCTGGAACGTG 1960
QY TGTAGCGAAGGCGGCTGTCGCGCGCGCGCTGTCGCGGAGGAG 1560
Db |||||
1961 GGTGGAGGAGGAGCGCGGAGTCTGCGCGGAGGAG 1998

```

RESULT 11

```

AF486515 1827 bp mRNA linear PLN 02-JUL-2002
LOCUS Hordeum vulgare cultivar waxy oderbrucker granule bound starch
DEFINITION synthase I mRNA, complete cds.
ACCESSION AF486515
VERSION AF486515.1 GI:21667429
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.
REFERENCE 1 (bases 1 to 1827)
AUTHORS Patron,N., Smith,A., Fahy,B., Hylton,C., Naldrett,M., Rossnagel,B.
and Denyer,K.
TITLE A mutation in the 5' non-coding region of the barley GBSI gene
alters its temporal and spatial expression and reduces GBSI
activity and amylose content in the endosperm
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1827)
AUTHORS Patron,N.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
FEATURES
source
1..1827
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="waxy oderbrucker"
/db_xref="taxon:4513"
1..1827
/note="alpha glucosyl transferase"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AA074049.1"
/db_xref="GI:21667430"
/translation="MAALATLSQATSGTVLGVTDTSAPSMFRHAGFQGLRPRKPADGT
FMRTVGAAPKQSKAHNRNRLSVVVRATGSMNLVFGAAMAPWPKYKRGVDR
LGGLDPAAMAANGHRVWVSPRYDQKADWTSVISEIKVADEVKRVFHCYKRGVDR
VFIDHPFLEKVRGKTKETIYGDAGTDEQDQRESLCCOALEPRLINNNPYF
SPYGEDVVCNDWHTGLIACLYKSNOSNGIYAKVAFCHINISYQGRSFSDFFA
QLNLPDRKSSDFDIDGDKPVEGRKINMKAGILOADKLVFVSPYAEELISGEARG
CELDNMLRIGITGVNGMDVSEMDPTKDFLAVNYDITALEAKALKEALQAEVGL
PVDKRPILVAFILRELEGKDPVMTAAIPEILKEEDVQILLEGKFKKLLKSMEE
KPKGVAVRFPNAPLAHOMAGADLLAVTSFPCGLIQLQGMRYGTPCVCASTGGL
VDTIVEGKTFHMRSLVDCNVPEADVKVATITLRAVKVVGTPAYQBMVKNCMIQD
LSWKGPAKNWEDVLELGVGSEPGIVGEETAPLAVENVAAP"
BASE COUNT 364 a 567 c 602 g 294 t
ORIGIN
Query Match 27.1%; Score 528.4; DB 8; Length 1827;
Best Local Similarity 61.6%; Pred. No. 7.2e-44;
Matches 984; Conservative 0; Mismatches 521; Indels 93; Gaps 5;
QY 5 TGGACATCGTGTGTTGCTGCTGAGTGCCTTGGTCCAGAGCGGCGCTGGCG 64
Db |||||
242 TGAACCTCGTGTTCGTCGCGCGCGAGATGGCGCTTGGAGCAAGACCGCGCTGGCG 301
QY 65 ATGTGACTGTGCTGCTGCTATTGAGTGTCTCAAGCGCGCGCACCGCTGACCATTTG 124
Db |||||
302 AGCTCTCGCGCGCTTCCAGCAGCATGGCGCGCACCGTACCGGGTGATGGTCTGCT 361
QY 125 CCCCTCGCTACGACAGTACGCTGACGCTGGGACACCTCGGTGCGTGGACATCATGG 184
Db |||||
362 CCCCGCGCTACGATCAGTACAGGACGCTGGACACCGAGCTCATCTCCGAGATCAGG 421
QY 185 -----GCCAGAGGTCCGCTACTTCCACTCCATCAAGAGCGCGTGCACCGCG 232
Db |||||
422 TCGCTGACGATACGAGAGGAGGTGAGGTTCTTCCACTGTACAAAGCGCGAGTGGACCGG 481
QY 233 TGTGATTGACCAACCCCTGGTTCCTGGCCCAAGTCTGGGCAAGACCGCTCCCAAGTGT 292
Db |||||
482 TGTTCATCGACCAACCCCTGGTTCCTGGAGAGGTTCCGGGCAAGACCAAGAGAGATCT 541

```



Qy	65	ATGTGACTGTGGCTTGCCTTATTTAGCTGTGTCAAGCGCGCCACCGGCTCATGACCATTGG	124
Db	302	ACGTCTCGCGGCGCTTCCACAGCCATGGCCGCCAACGGTCACCGGGTGAATGGTCTGCT	361
Qy	125	CCCTTCGCTACGACCAAGTAGCGCTTGAGCGCTTGGGACACCTTCGGTGTGCTGGACATCATGG	184
Db	362	CCCGCGCTACGATCAGTACAAAGAGCGCTTGGGACACCAAGCGTCACTCCGAGATCAAGG	421
Qy	185	-----GGGAGAAGTCCGCTACTTCCACTTCCATCAAGAAGGCGGTGCAACCGG	232
Db	422	TCGCTGACGAGTAGACGAGGGTGGAGTCTTCCACTGCTACAAAGGCGGAGCTGGACCGG	481
Qy	233	TGTGATTGACCAACCCCTGGTCTTCCGCAAGGCTTGGGGCAAGACCGGCTCCAAAGCTGT	292
Db	482	TGTTTCATCGACCAACCCGTGGTTCCTGGAGAGGTCGCGGGCAAGACCAAGAGAAGATCT	541
Qy	293	ACGGCCCCCGCTCCGGGCTTGACTCTGGACAAACCAAGCGCTTCGCGCTGTGTTCTGCA	352
Db	542	ACGGCCCCGACGCGGCACGGACTACGAGGACAACCAAGCAGCGCTTCAGCGCTTCTCTGCC	601
Qy	353	AGGCGCGTATTGAGGCTGCCGGTCTGCCCTTC-----G	388
Db	602	AGGACGGCTCGAGGACACCCAGGATCCTCAACTCAACAACACCCCTACTTTCTGTGC	661
Qy	389	GCCCCGCGAGGACTCGCTTCCTGGTGGCAAGACTTGGCACTCCGCGCTGTGCGCGCTCC	448
Db	662	CCTACGGGAAGACGTGGTTCGTGTGCAACGACTGGCACAGGGCCCTTTCGGCCTGCT	721
Qy	449	TGCTGAAGGACGAGTACCAAGCCCAAGGGCCAGTTCAACAAGGCCAAGTCGTTGCTGCGCTA	508
Db	722	ACCTCAAGAGCAACTACCAAGTCCATATGGCATCTACAGACGGCCAAGGTGGCCCTTCTGCA	781
Qy	509	TCACAACATCGCTTTCAGAGGCGCATGTGGAGAGGCTTTCAGGACACAGAGCTGC	568
Db	782	TCACAACATCTGTAACAGGGCGGCTTCTCTCGACGACTTTGCGCAGCTCAACCTGC	841
Qy	569	CCAGGCGCGCTTTGACAAGCTGGGCTTCTCGAGCGGCTATGCCAAGGTTTACACTGAGG	628
Db	842	CCGACAGGTTCAAGTCGTCTTCGACTTCATTGACGGCTACGACAG-----888	
Qy	629	CCACCCCATGGAGGACGAGAGACGCCCGCTGACGGGAAGACCTTACAAGAAGATCA	688
Db	889	-----CCCGTGGAGGGGCG-----AAGATCA	910
Qy	689	ACTGGCTGAAGGTTGGCATTTACGCGCGCAAGCTGTGACTGTGTGCGCCCACTACG	748
Db	911	ACTGATGAAGCCGGATCTCTGAGGCGGACAAAGTCTGACGGTAGCCCTTACTACG	970
Qy	749	CGACCGAGATTCGCTGCCGATGCCCGGCGGTGTGGAGCTGGACACGCTCATCCGCGCCA	808
Db	971	CTGAGAGCTCATCTCTGGCGAAGCAGGGGCTGGAGCTCGACAACATCATGCGGCTCA	1030
Qy	809	AGGGCATTCAGGGCATTTGAACGCGCATGGACATTGAGAGTGGAAACCCCAAGACCGACA	868
Db	1031	CGGGATCACCGGAATCGTCAACGCGCATGGAGCTGAGTGGAGGACCCCAAGAGACA	1090
Qy	869	AGTTCCTGTCTGCGCCCTACGACCAAGACAGCGTCTACGCGCGCAAGCGCCGCCAAGG	928
Db	1091	AGTTCCTCGCGCTCACTACGACATACCAACCGGCTTGGAGGCCAAGGCGTGAACAAAGG	1150
Qy	929	AGGCCCTCGAGCGGAGCTGGGCGCTGCTGTGGACCCACCGCCCGCTGTTCGCCCTCA	988
Db	1151	AGGCACTCGAGCCGAGTGGGGCTGCCGTTGGACAGGATGCGCGCTGCTGGCTTCA	1210
Qy	989	TCGGCGCCTTGAGGACGAGAGGCTGTGGACATCATCTCTGGCGCGCTGCCCAAGATCC	1048
Db	1211	TCGGCAGGCTGGAGGACGAGAGGGCCCGACGTGATCGCCCGCATCCCGGAGATCC	1270
Qy	1049	TGCCCCACCCCAAGTTCGAGATCGCCATCTCTGGTACCGGCAAGGCCCGCTACGAGAAGC	1108
Db	1271	TGAGGAGGAGACGTCCAGATCATTTCTTGGCACCGGAGAGAGATTTGAGAAGC	1330

QY	1109	TGGTGAACCCCATCGCACCAAGTACAAAGGGCGGCCAAAGGCGTGGTCAAAATTCCTGG	1168
Db	1331	TGCTCAAGAGCATGGAGGAGAAGTTCCTCCGCGAAGGTGAGGCGCGTGGTCAGGTTCAACG	1390
QY	1169	CGCCCTGTGCGCACATGTCACCGCGGCGCGACTTCATGCTGGTGCCCTCGCGCTTCG	1228
Db	1391	CGCCGTAGCTACCCAGATGATGGCGGCGCGCACTTGCTGCTGTACACCGCGCTTCG	1450
QY	1229	AGCCTGTGGGCTGATCCAGCTGCACGCCATACACTACGGTACCGGTGCCCGTGGTAGCCT	1288
Db	1451	AGCCCTGTGGGCTCATCCAGCTCCAGGAAATCGCTATGGAACGCCGTGCGTGTGCGGT	1510
QY	1289	CCACCGGGGCTGTGTCACACCGTCAAGGAGGGCGTCACCGGCTTCACATGGGGGCC	1348
Db	1511	CCACCGGGGCTGTGTCACACGATCGTGAGGGCAAGACCTGGTTCCACATGGGCGCG	1570
QY	1349	TGA-----ACCCGACAAAGCTGGACGAGGCTGAGCGCGACGCCCTGGCGCCACCGTGC	1402
Db	1571	TACAGCTCGACTGCAACGTGTGTGAGCGCGGACGTGAAGAAGGTGGCGACCACTTGA	1630
QY	1403	GCCGTGCACGAGATGTGTCGGGGGCGCGGTACCCGAGATGTGGCGCAACTGCATCA	1462
Db	1631	ACGGGGCGTCAAGGTCTGTCGCGACCGCGGGGTACGACGAGATGGTCAAGAACTCATGA	1690
QY	1463	GCCAGGACTGTCTCTGTGTCCAAGCCCGCCCAAGCTGGAGGGCGTCTGTGGAGAGGTGG	1522
Db	1691	TCCAGGATCTCTCTGTGAAGGACCTGCCAAGAACTGGAGGACGTGTTCTGGAACTGG	1750
QY	1523	TGTACGGCAAGGGCGGCTGGGCCACCGCCAAAGAGG	1560
Db	1751	GGGTGGAGGGAGAGCGCGGGGATCGTCTGGCGAGGAG	1788
RESULT	13		
AF163319			
LOCUS		2127 bp	linear
DEFINITION		Triticum aestivum granule-bound starch synthase I (Wx-D1) mRNA, Wx-D1a allele, complete cds.	PLN 14-FEB-2000

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

```

/protein\_id="AAF34135.1"  
/db\_xref="GI:696978"  
translation="MAALVTSQLATSGTVLGTIDRRFRAGFGVRRPRPADAAALGMRT  
VGASAAPTQSKAHRTGTRCLSMVVRATSGSGMNLVFGAEMAPWSKTGGLDVLGGL  
PPMAANGHRVMSIPRIDQTKDAMDTSVSEIKVVDYKIERVRYPHCHYKRGVDYFVYD  
HPCFLKVRGTRKIKYIGPDAGTDYEDNQRFSLCCQAALVPRILINDNNPDRFSGPY  
GEDVYFVNDHTGLLACYLKSNSNGIYRAAKVAFCHINI SYGRFSFDDPAQLNL  
PDRFKSSPFDIDGDKPVEGRKINNMKAGILQADKLVSPYIAELISFGEARGCELD  
NIMRLTGTGVNGMDVSEMDPTKDFELAVNDITTALEGRALNKEALQAEVGLPVDLR  
KVLVAFIIGRLLEOKGPDYMTAAIPEILKEEDVOILVLTGCKKFERLLKSTEEKFP  
KYRVRRENAELHOMAGADVLAIVSRFEPGCLLQLOGMRYTGPCACASUGGLVDIT  
VEGTFKFMHGRSLSDNVYEPADVKVKTITLKRKVVGTTPAYHEMVNKMCIQDLISWK  
GPAKNMVDLLELGVGSEPGVIGIEIAPLAMENVAAP"  
BASE COUNT 452 a 617 c 687 g 371 t  
ORIGIN

Query Match 26.9%; Score 525.6; DB 8; Length 2127;  
Best Local Similarity 61.4%; Pred. No. 1.3e-43;  
Matches 1002; Conservative 0; Mismatches 534; Indels 96; Gaps 6;  
QY 5 TGGACATCGTGTGCTGTGCTGAGTCCGCCCTTGGTCCAAAGACGGCGGCTGGGG 64  
DB 259 TGAACCTCGTGTTCGTCGGCGCGAGATGGCCCTTGGACAGACGGCGGCTGGCG 318  
QY 65 ATGTGACTGTGGCCCTGCTATTGAGTGTGTCAAGCGCGGCCACCGGCTGATGACCATG 124  
DB 319 ACGTCTCGGGGCTCCGCCAGCCATGCGCGCCCAAGCCACCGGCTCATGTCTATCT 378  
QY 125 CCCCTCGCTACGACAGTACGTGACGCTGGGACACCTCGTGTGCTGGGACATATGG 184  
DB 379 CCCCAGCTACGACAGTACGAGGACGCTGGGACACACAGCTGCTCGGAGATCAAGG 438  
QY 185 -----CGGAGAAGTCCGCTACTTCCACATCCATCAAGAAGGGCGCTCACCGCG 232  
DB 439 TCGTTCACAGTACGAGAGGTGAGGTACTTCCACTGCTACAGCGGGGTGACCGCG 498  
QY 233 TGTGAGTACACACCGCTGTCTCTGGCCAAAGTCTGGGGCAAGACCGGCTCCAAAGTGT 292  
DB 499 TGTCTGACACACCGCTGTCTCTGGAGAAAGTCCGGGCAAGACCAAGGAGAATCT 558  
QY 293 ACGGCCCCGCTCCGGCGTACTACTGGACACACACAGCGCTCGCCCTGTCTGCA 352  
DB 559 ACGGGCGCGACCGCGCGAGGATAGGAGGACACACAGCGCTTACGCTTCTCTGCC 618  
QY 353 AGCGCGCTATTGAGGTGCGCGGCTGCTGCCCTTCG----- 388  
DB 619 AGCGCGCTGGAAGTCCGAGGATCCTGAACCTCGACATAACCCCTACTTTCTTGGCG 678  
QY 389 GCGCGCGGAGGACTGCTGCTGCGCCACAGCTGGGCACTCGCGCTGGTGGCGCTCC 448  
DB 679 CCTACGGGAGGAGCTGGTGTCTGCAATGATGATGATGATGATGATGATGATGATGAT 738  
QY 449 TGTGAGGAGGAGTACACAGCCCAAGCGGCTTCAAGAGGCAAGTGGTGTGCTGCTA 508  
DB 739 ACCTAGAGCACTACCACTCAATGGCATCTACAGGCGCGCAAGGTCGATCTGCA 798  
QY 509 TCCACACATCGCTTCCAGGCGGATGTTGGAGGAGGCTTTTCAAGGACACGAAAGTGC 568  
DB 799 TCCACACATCTCGTACAGGCGGCTTCTCTTCGACGACTTCGCGCAGCTCAACCTGC 858  
QY 569 CCAGGCGCGCTTGTGACAGCTGCGCTTCTCGGACGCTATGCCAAGTTTACACTGAGG 628  
DB 859 CCAGAGGTTCAAGTGTCTCTCGATTCATCGACGCTAGACAAAG----- 905  
QY 629 CCACCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 688  
DB 906 -----CGGTTGAGGGGCGC----- 927  
QY 689 ACTGGCTGAGAGGTGGCATTTATCGCCCGCGGACAGCTGTTGCTGCTCGCCCAACTACG 748  
DB 928 ACTGATGAGAGGCGGAGTCTCTGAGGCGGACAGAGGTCGTGAGGCTGAGGCTGAGGCT 987  
QY 749 CGACCGAGATCGTGTCCGATCCCGCGGCTGTGTGGAGTGGACACCGCTCATCGCGCCA 808

DB 988 CGGAGAGCTCATCTCTGGCGAAGCAGGGGCTGCGAGCTCGACAAACATCATCGGCTCA 1047  
QY 809 AGGGCAITGAGGGCATTTGTGAACGGCATGACATTTAGGAGTGGAAACCCCAAGACCGACA 868  
DB 1048 CTGGGATCACCGGCATCTGCAACGGCATGATGTTAGCACTGGACCCCAAGGACA 1107  
QY 869 AGTTCTGTCTGCGCCCTACGACCAAGAACAGCTCTACGCCGGAAGCGCGCGCAAGG 928  
DB 1108 AGTTCTGCGCGTCAACTACGACATCACACCGCTTGGAGGGAAGGCGCTGAACAAG 1167  
QY 929 AGCCCTCGAGGCGAGCTGGGCTGCTGTGACCCCAAGCGCGCGCTGTTGCGCTCA 988  
DB 1168 AGCGCTGACGCGGAGTGGGCTGCGGTGACCGGAAGTGCCTGTTGGCGGTCA 1227  
QY 989 TCGCGCGCTTGGAGGAGCAAGGCTGTGACATCATCTTGGCGCGCTCCCAAGATCC 1048  
DB 1228 TCGGAGGCTGGAGAGCAAGAGGCGCGCGCTGATGATGCGCGCATCCGAGATCC 1287  
QY 1049 TGCCCAACCCCAAGGTGACATCGCCATCTGTGGTACCGGCAAGCGCGCTACGAGAC 1108  
DB 1288 TGAAGGAGGAGGAGCTCCAGATCTTCTCTGGGCAACCGGAAGAGTTCGAGCGC 1347  
QY 1109 TGTGAGCGCATCGGCACCAAGTACAAGGCGCGGCGCCCAAGGGGTGTCAAGTTCG 1168  
DB 1348 TACTCAAGAGCATTTGAGGAGAAATCCCGAGCAAGTGTGAGGCGCTGTTGAGTCAACG 1407  
QY 1169 CGCCCTTGGCGCATGCTCACCGCGCGCGCTTCTGCTGTGTTGCTCGCGCTTCG 1228  
DB 1408 CGCGCTGGCTCACCAGATGATGCGCGCGCGCTGCTGCTGCTCACCGCGCTTCG 1467  
QY 1229 AGCCCTGCGGCTGATCCAGCTGCAACCGCATGCACTACGCTACGCTGCTGCTGCTG 1288  
DB 1468 AGCCCTGCGGCTCATCCAGCTCCAGGGGATGCGCTACGGAAGCGCGCTGCGCGCT 1527  
QY 1289 CCACCGCGGCTTGGTGTGACACCGCTCAAGAGGCGCTCACCGCTTCCACATGGCGCC 1348  
DB 1528 CCACCGCGGCTTGTGACACGATCTGGAGGCAAGACCGGCTTCCATGAGGCGCGC 1587  
QY 1349 TGA-----ACCCCGCAAGCTGGACAGGCTGACCGCGCGCTGCGCGCTGCGCGCTG 1402  
DB 1588 TCAGTGTGATTCAGAGTGTGAGCGCGCGCGCTGAGAGAGTGTGTGACCACTG 1647  
QY 1403 GCGGTGCGAGGAGTGTGCGGCGCGCGCTACCGGAGAGTGTGCGCACTGCA 1462  
DB 1648 AGCGCGGCTCAAGCTGCTGCGCACCGCGCATACATGAGATGTTCAAGAACTGCA 1707  
QY 1463 GCCAGGACCTGCTCTGTTCCAAAGCGCGCGCAAGTGGAGGCGCTGCTGAGGAGTGG 1522  
DB 1708 TACAGATCTCTCTGAAAGGCGCGCGCAAGAACTGGGAGAGTGTCTTGAAGTGG 1767  
QY 1523 TGTACGCAAGGCGGCTGCGCCACCGCAAGAGAGGAGATCAAGTGTGCGCTGCGCG 1582  
DB 1768 ---GTGTCGAGGAGGAGCGCGGCGGCTCATCGGCGAGGATGCGCGCTGCCATGG 1824  
QY 1583 AGAATATCCCG 1594  
DB 1825 AGAAGTTCGCG 1836

RESULT 14  
AF486519  
LOCUS  
DEFINITION  
AF486519  
ACCSSION  
AF486519  
VERSION  
AF486519.1  
KEYWORDS  
SOURCE  
ORGANISM  
Hordeum vulgare.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
AF486519  
Hordeum vulgare cultivar CDC Alamo granule bound starch synthase 1  
mRNA, complete cds.  
1812 bp mRNA linear PLN 02-JUL-2002  
GI:21667437









Db 1691 TCCAGGATCTCTCCTGGAAGGACCTGCCAAGAACTGGGAGGACCTGCTTCIGGAACTGG 1750

QY 1523 TGTRCGGCAAGGGCGCGTGGCCACCGCCCAAGAGGAG 1560

Db 1751 GGTGGAGGAGCGAGCCGGGATCTCTCGGCGAGGAG 1788

Search completed: June 3, 2003, 13:34:07  
Job time : 3422.54 secs

**This Page Blank (uspto)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 05:08:24 ; Search time 283.834 Seconds  
(without alignments)  
15495.523 Million cell updates/sec

Title: US-09-980-771A-4

Perfect score: 1953

Sequence: 1 gcgctgacatcgtatggt.....ccgccaccacgaagtcgcgc 1953

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	704	36.0	1696	22	AAAC86955
2	475.8	24.4	2267	18	AAAC63355
3	471	24.1	1915	19	AAAC60319
4	447.4	22.9	2542	19	AAV29753
5	316.8	16.2	2161	24	ABK53210
6	298	15.3	2115	21	AAAC32824
7	258	13.2	2183	24	ABK53215
8	221	11.3	2807	21	AAAC86435
9	221	11.3	2842	21	AAAC86411

10	218.2	11.2	2826	19	AAV01528
11	216.8	11.1	2097	19	AAV29755
12	214.8	11.0	2423	19	AAV70958
13	214	11.0	2107	21	AAAC86412
14	214	11.0	2107	21	AAAC86434
15	212.4	10.9	9024	21	AAAC86431
16	210.8	10.8	2380	19	AAV56834
17	210.2	10.8	2920	24	ABK88112
18	209.8	10.7	2478	18	AAAT67287
19	209.2	10.7	2939	21	AAAC86410
20	208.6	10.7	2939	21	AAAC86433
21	208.6	10.7	2951	24	ABK88114
22	208.2	10.7	2950	24	ABK88113
23	208.2	10.7	2950	24	ABK88113
24	207	10.6	2248	21	AAZ50651
25	206.6	10.6	1798	21	AAZ50647
26	206.6	10.6	1798	21	AAZ50646
27	203.8	10.4	1724	24	ABA01821
28	161	8.2	2007	19	AAV70959
29	161	8.2	2007	19	AAV29754
30	161	8.2	2085	18	AAAT67286
31	161	8.2	2085	19	AAV66833
32	151	7.7	1479	24	ABQ90258
33	151	7.7	1479	24	ABQ90406
34	125.6	6.4	333	21	AACT4569
35	125.6	6.4	333	24	ABN18316
36	118.2	6.1	2239	19	AAV01527
37	118.2	6.1	2805	21	AAZ24487
38	116	5.9	2662	20	AAK34651
39	110.8	5.7	65140	22	AAI17184
40	110.8	5.7	125401	22	AAI17186
41	110.6	5.7	2533	15	AAQ45183
42	108	5.5	1758	17	AAAT32325
43	106.8	5.5	2277	19	AAV13836
44	106.8	5.5	2277	19	AAV05372
45	106.4	5.4	1984	21	AAAC46797

#### ALIGNMENTS

#### RESULT 1

AAAC86955	AAAC86955 standard; cDNA; 1696 BP.
ID	AAAC86955
XX	AC
XX	AC
XX	02-APR-2001 (first entry)
DT	02-APR-2001 (first entry)
XX	CDNA encoding a granule bound starch synthetase II (GBSSII).
DE	Granule bound starch synthetase II; GBSSII; starch grain;
XX	adenosine diphosphate glucose-alpha1,4-glucan alpha4-glucosyltransferase;
KW	starch synthetase; ds.
XX	Chlamydomonas reinhardtii.
OS	Chlamydomonas reinhardtii.
XX	Key
FH	Location/Qualifiers
CDS	3..719
FT	/*tag= a
FT	/product= "granule bound starch synthetase II (GBSSII)"
XX	FR2793806-Al.
PN	24-NOV-2000.
XX	21-MAY-1999; 99FR-0006494.
PD	21-MAY-1999; 99FR-0006494.
XX	21-MAY-1999; 99FR-0006494.
XX	(CNRS ) CNRS CENT NAT RECH SCI.
XX	Dhulst C, Ball S;
PI	

Wheat granule-bound  
Zea mays soluble s  
DNA encoding maize  
Wheat starch synth  
Wheat starch synth  
Wheat SSII gene SE  
Zea mays soluble s  
cDNA encoding modi  
Soluble starch syn  
Wheat starch synth  
Wheat starch synth  
cDNA encoding barl  
cDNA encoding barl  
cDNA encoding barl  
Corn starch syntha  
Corn starch syntha  
Rice starch syntha  
DNA encoding maize  
Zea mays soluble s  
Soluble starch syn  
Zea mays soluble s  
M. capsulatus gene  
Human ORFX ORF124  
Human ORFX polynuc  
Wheat soluble star  
Wheat soluble star  
cDNA sequence of w  
Streptomyces nous  
Streptomyces nous  
Soluble rice starch  
Soluble starch syn  
Homo sapiens mamma  
Human telomerase p  
Arabidopsis thalia

```
XX WPI; 2001-052291/07.
DR P-PSDB; AAB31175.
XX
PT New recombinant nucleic acid encoding fusion of starch synthase and
PT second component, useful in pharmaceutical and food compositions, is
PT targeted to starch granules -
XX
XX Claim 3; Fig 1; 52pp; French.
XX
XX The present sequence encodes a granule bound starch synthetase II
XX (GBSII). The specification describes a recombinant nucleic acid,
XX comprising a polynucleotide encoding an adenosine diphosphate
XX glucose-alpha1,4-glucan alphasyltransferase or starch
XX synthetase, placed upstream of a sequence that encodes a polypeptide
XX of interest. The adenosine diphosphate glucose-alpha1,4-glucan
XX alphasyltransferase protein can migrate to sites of
XX biosynthesis of starch grains in plant cells, becoming associated
XX with these grains. The recombinant nucleic acid sequence is used to
XX target polypeptides of interest to starch grains.
XX
XX Sequence 1696 BP; 347 A; 509 C; 571 G; 269 T; 0 other;
SQ
Query Match 36.0%; Score 704; DB 22; Length 1696;
Best Local Similarity 100.0%; Pred. No. 7.6e-93;
Matches 704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1250 TGCACGCCATGCACTAGCTACGTCGCCGTGTAGCTCCACCGCGCTGGTGCACA 1309
DB 13 TGCACGCCATGCACTAGCTACGTCGCCGTGTAGCTCCACCGCGCTGGTGCACA 72
QY 1310 CCCTCAAGGAGGCGTCACCGGCTTCCACATGGCGCCCTGACCCCGACAAGCTGACG 1369
DB 73 CCCTCAAGGAGGCGTCACCGGCTTCCACATGGCGCCCTGACCCCGACAAGCTGACG 132
QY 1370 AGCTGACGCGCAGCGCCCTGCGCGCCGACCGCTGCGCGCTGCCAGGAGTGTTCGCGGCG 1429
DB 133 AGCTGACGCGCAGCGCCCTGCGCGCCGACCGCTGCGCGCTGCCAGGAGTGTTCGCGGCG 192
QY 1430 GCCGCTACCCCGAGAGTGGTGGCCAACTGCATCAGCCAGGACCTGCTCTGTCTCAAGCCG 1489
DB 193 GCCGCTACCCCGAGAGTGGTGGCCAACTGCATCAGCCAGGACCTGCTCTGTCTCAAGCCG 252
QY 1490 CCAGAGTGGAGGCGCTCTGTGAGGAGTGTGTACGCAAGGGCGGCTGGCCACCG 1549
DB 253 CCAGAGTGGAGGCGCTCTGTGAGGAGTGTGTACGCAAGGGCGGCTGGCCACCG 312
QY 1550 CCAGAGGAGGAGATCAAGTCCCGTTCGCCAGAGATCCCGGGGACCTGCGCGCG 1609
DB 313 CCAGAGGAGGAGATCAAGTCCCGTTCGCCAGAGATCCCGGGGACCTGCGCGCG 372
QY 1610 TGTCTACGCCCCCAACACCCCTGAAGCCGCTGTCGCGCTCCGTGGAGGGCAAGCGCGCG 1669
DB 373 TGTCTACGCCCCCAACACCCCTGAAGCCGCTGTCGCGCTCCGTGGAGGGCAAGCGCGCG 432
QY 1670 CCAGCGCCCAAGGTGGCGACACCGCCCGCCCGATGGCGCGTGGCGCGACCAACCCCT 1729
DB 433 CCAGCGCCCAAGGTGGCGACACCGCCCGCCCGATGGCGCGTGGCGCGACCAACCCCT 492
QY 1730 CGGCGCCCTGCGCGCGCGCGCCACCCCAAGGTGACCACTACAGCCCGCCCTGCGCG 1789
DB 493 CGGCGCCCTGCGCGCGCGCGCCACCCCAAGGTGACCACTACAGCCCGCCCTGCGCG 552
QY 1790 CCAGCGCCCAAGCGCCCTGCGCGCTCAAGCTGGCGCGTGGCGCGTCCACCACTCGA 1849
DB 553 CCAGCGCCCAAGCGCCCTGCGCGCTCAAGCTGGCGCGTGGCGCGTCCACCACTCGA 612
QY 1850 CCTCGGAGAACGGCGCTGCTTCCAAACGGCAACGGCAACGGTGGCTCGGCTTCCAAAGACCT 1909
DB 613 CCTCGGAGAACGGCGCTGCTTCCAAACGGCAACGGCAACGGTGGCTCGGCTTCCAAAGACCT 672
QY 1910 CGGCTGACAGCCCTGCTCGCGCGCGCCACCGCGAGTCCGCC 1953
|||||
```

DB 673 CGGCTGCCAAGCCCTGGTCTCGCGCGCCACCGCAAGTCCGCC 716

#### RESULT 2

```
AXX63355
ID AAX63355 standard; cDNA; 2267 BP.
XX
AC AAX63355;
XX
XX 16-JUL-1999 (first entry)
XX
XX Granule bound starch synthase encoding cDNA.
DE
XX Maize; corn; Zea mays; Delta-9 desaturase; GBSS; target; substrate;
KW granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;
KW modulation; gene expression; transgenic plant; cleavage; canola plant;
KW caffeine synthesis; coffee plant; nicotine production; tobacco;
KW fruit ripening; flower pigmentation; lignin production; ss.
XX
XX Zea mays.
XX
XX WO9710328-A2.
XX
XX 20-MAR-1997.
XX
XX 12-JUL-1996; 96WO-US11689.
XX
XX 13-JUL-1995; 95US-0001135.
XX
XX (DOWC ) DOWELANCO.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX
XX Edington BE, Folkerts O, Guo L, McSwiggen JA, Merlo DJ;
PI Merlo PAO, Skokut TA, Young SA, Zwick MG;
XX
XX WPI; 1997-202224/18.
XX
XX Ribozyme which modulates plant gene expression - preferably
PT modulates expression of DELTA-9 desaturase or granule bound starch
PT synthase in maize or canola
XX
XX Example 9; Page 31-33; 155pp; English.
XX
XX The present invention describes an enzymatic nucleic acid molecule (I)
CC with RNA cleaving activity, which modulates the expression of a plant
CC gene. Also described is a gene comprising a cDNA sequence encoding maize
CC Delta-9 desaturase. (I) can be used to modulate expression of a gene,
CC preferably Delta-9 desaturase or a granule bound starch synthase (GBSS)
CC gene, in a plant (preferably a maize or canola plant). (I) can be used
CC to modulate caffeine synthesis in a coffee plant, nicotine production in
CC a tobacco plant, fruit ripening processes in an apple, tomato, pear,
CC plum or peach plant, flower pigmentation in a rose, petunia,
CC chrysanthemum or marigold plant or lignin production in a tobacco,
CC aspen, poplar or pine plant.
XX
XX Sequence 2267 BP; 465 A; 699 C; 733 G; 370 T; 0 other;
SQ
Query Match 24.4%; Score 475.8; DB 18; Length 2267;
Best Local Similarity 59.9%; Pred. No. 4.1e-60;
Matches 979; Conservative 0; Mismatches 557; Indels 99; Gaps 7;
QY 5 TGGACATCGTGATGGTGTGCTGAGTGGCCCTTGGTCCAAAGCGGGCGCTGGCG 64
DB 418 TGAACGTCGCTTTCGTGGCGCCGAGATGGCGCGTGGAGCAAGACCGGGCGCTCGCG 477
QY 65 ATGTGACTGTGTGGCTGCTATTGAGCTCAAGCGCGCGCTCATGACCAATTG 124
DB 478 ACGTTCCTCGCGGCTGCCCGCGCCATGGCCGGAATGGCACCGTCTCATGTGCTCT 537
QY 125 CCCCTCGCTACGACCACTACGCTACGCGCTGGGACACCTCGGTGGTCTGTGACATF 179
DB 538 CTCGCCGCTACGACCACTACGCTACGCGCTGGGACACCGCTGGGAGTCTGTCTCCGAGATCAAGA 597
|||||
```

QY 180 -----CATGGGAGAGTCCGCTACTTCCATCTCATCAAGAGGGGGTGCACCGCG 232  
DB 598 TGGAGACAGGTACGAGACGCTCAGGTTCTTCCACTCTCAAGCGCGAGTGCACCGCG 657  
QY 233 TGTGGATTGACCAACCCCTGCTTCTGCGCAAGTCTGGGCAAGACCGGCTCCAAGCTGT 292  
DB 658 TGTTCGTTGACCAACCCCTGCTTCTGAGAGGGTTTGGGAAAGACAGAGAGAATCT 717  
QY 293 ACGGCCCCCGCTCCGGGCTGACTACTGTCGACAAACAGCGCTTGGCCCTGTCTTCTGCA 352  
DB 718 ACGGGCTTGACGCTGGAACGAGCTACAGGACAAACAGCTCGGTTTCAGCTGCTATGCC 777  
QY 353 AGGCGCTATTGAGGCTGCCCGGTGCTGCCCTTCGGC----- 390  
DB 778 AGCAGCATTGAGCTTCAAGATCTGAGCCTCAACAACACCCCAIATCTTCCGGAC 837  
QY 391 --CCCGGCGAGGACTGGCTTCTGTCGCCAACGACTGGCACTCGCCCTGTGTCGCCGTC 448  
DB 838 CATACGGGAGGACGCTGCTTCTGCTGCAACGACTGGCACTCGCCCTCTCTGCTGT 897  
QY 449 TGTGAGGACGAGTACCAAGCCCAAGGGCCAGTTTACCAAGGCGCAAGTCTGGTCTGCTA 508  
DB 898 ACCTCAAGAGCACTACCACTGCTCCAGTCCAGGCACTTACAGGACGCAAGACCGCTTCTGCA 957  
QY 509 TCCACAACATCGCTTCCAGGGCGCATGTGGGAGGAGGCTTCAAGGACACAGACTGC 568  
DB 958 TCCACAACATCTCTACCAAGGCGGTTGCGCTTCTCCGACTACCCGAGCTGAACCTOC 1017  
QY 569 CCCAGGCGCTTTGACAGCTGGCTTCTCGACGCGCTATGCCAAGGTTTACACTGAGG 628  
DB 1018 CGGAGAGATTCAAGTCTGCTTCTGATTTATCGACGCTACGAGAAG----- 1064  
QY 629 CCACCCCATGAGGAGGAGGACGAGAACCCCGCTGACGGGAAGACCTACAAGAATCA 688  
DB 1065 -----CCCGTGGAG-----GCCGGAAGATCA 1086  
QY 689 ACTGCTGAAGGTGGCTATTATCGCCGCGACAGAGTGGTACTGTCTGCCCAACTACG 748  
DB 1087 ACTGGATGAAGCGCGGATCTCTGGAGCGGACAGGTTCTACCGTACGCCCCCTACTACG 1146  
QY 749 CGACCGAGATCGTCCGATGCGCGCGGGTGTGAGCTGGACACCGTATCCGCGCCA 808  
DB 1147 CCGAGAGATCTATCTCCGCTATCGCAGGCGGTGCGAGCTCGACACATCATCGGCTCA 1206  
QY 809 AGGCGATTGAGGCGATTGTACGGGATGACATTTAGAGAGTGAACCCCAAGACGACA 868  
DB 1207 CCGGATACCGGATCTGCAAGCGGATGAGCTGACGAGTGGGACCCCGCAGCAGGACA 1266  
QY 869 AGTTCTGTCTGGCCCTACGACAGACAGCTCTACGCGGCAAGCGCGCGCAAGG 928  
DB 1267 AGTACATCGCGTGAAGTACGAGCTGTGACGCGCGCTGGAGGCAAGCGCTGAACAAG 1326  
QY 929 AGGCGCTGACGCGGAGCTGGGCTGCTGTGACCCCAAGCGCCCTGTTGGCCCTTCA 988  
DB 1327 AGGCGCTGACGCGGAGCTGGGCTGCTGTGACCGGAACATCCGCTGTTGGCGCTTCA 1386  
QY 989 TCGGCGCTGAGGAGGAGAGGAGTGTGACATCATCTGCGCGCTGCCCCCTGCCCCATCC 1048  
DB 1387 TCGGAGGCTGGAAGAGCAGAGAGGACCCGACGCTACGCGCGCGCCATCCCGCACTCA 1446  
QY 1049 TGGCCACCCCAAG---GTGCAATCGCCATCTGCTGAGTACCGGCAAGCGCGCTACGAGA 1105  
DB 1447 TGGAGATGTGGAGGAGCTGAGATGTTCTGTGCGACGCGGCAAGAGAGATTTCGAGC 1506  
QY 1106 AGCTGTGAACCCATCGGCACCAAGTACAGGGCGCGGCCCAAGGGCGTGGTCAAGTTCT 1165  
DB 1507 GCGATGCTATGAGCGCGGAGAGAGATTCCAGGAAGGTGCGCGCTGGTCAAGTTCA 1566  
QY 1166 CGGCGCCCTGCGGACATGCTCACCGCGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCT 1225  
DB 1567 ACGGCGGCTGCGGACCACTATGCGCGCGCGGACGCTGCTGCTGCTGCTGCTGCTGCT 1626  
QY 1226 TCGAGCCCTGCGGCTGATCCAGCTGACGCGCATGCTACGCTACGCTGCGCGTGGTAG 1285

DB 1627 TCGAGCCCTGCGGCTCATCTCAGCTGCAGGGGATCGGATACGGAAGCCCTGCGCTGCG 1686  
QY 1286 CTTCCACCGGGGCTGCTGTCACACCTCAAGAGAGGGGCTCACGGCTTCCACATGGGCG 1345  
DB 1687 CGTCCACCGGGTGGACTGCTGACACCATCATCTCAAGGCAAGACGGGTTCCACATGGGCC 1746  
QY 1346 CCCTGA-----ACCCCGACAAAGCTGGACGAGGTGACGGCGACGCCCTGCGCGCCACCG 1399  
DB 1747 GCCTCAGCTGCACTGCAACGCTGCTGAGCGCGGAGCTCAAGAGTGGCCACCACT 1806  
QY 1400 TGCCTGTCAGGAGGCTGTTTTCGGGCGCGCTACCCCGAGATGTCGCAACTGCA 1459  
DB 1807 TGCAGCGCGCATCAAGGTGCTGCGCACCGCGGCTACGAGGAGATGTTGAGGAAGTCA 1866  
QY 1460 TCAGCCAGGACCTGCTGCTGCTCAAGCGCGCCCAAGTGGGAGGCGCTGCTGAGGAGG 1519  
DB 1867 TGATCCAGATCTCTCTGGAAGGGCTTCCAAAGACTGGGAGAACTGCTGCTCAGCC 1926  
QY 1520 TGTGTTACGCAAGGCGGCTGCGCCACCGCAAGAGAGGAGATCAAGTGCCTGTTG 1579  
DB 1927 TCGGGTTCGCC---GGCGGAGCGCAGGCTGCAAGGCGAGGATCGCGCGCTCGCCA 1983  
QY 1580 CCGAGAAGATCCCGG 1594  
DB 1984 AGGAGAAGTGGCGG 1998

## RESULT 3

AX60319  
ID AX60319 standard; DNA; 1915 BP.  
XX  
AC AX60319;  
DT 23-AUG-1999 (first entry)  
XX  
DE DNA sequence of the maize waxy gene.

XX Non-glycogen-like polysaccharide production; fermentation; waxy gene;  
KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;  
KW non-starch branching gene; amylopectin; amylose; plant-like starch; ss.  
XX

OS Zea mays.

XX W09844780-AL.

XX 15-OCT-1998.

XX 03-APR-1998; 98WO-US06660.

XX 04-APR-1997; 97US-0042939.

XX (EXSE-) EXSEED GENETICS LLC.

XX Guan H, Keeling PL;

XX WPI; 1998-568285/48.

XX P-PSDB; AAY16604.

XX Producing non-glycogen-like polysaccharides in bacteria, fungi or  
PT plants - transformed with genes for enzymes involved in starch or  
PT glycogen synthesis allows fermentative production of starches with  
PT engineered properties

XX Disclosure; Fig 49; 150pp; English.

XX The specification describes a method for the production of  
CC non-glycogen-like polysaccharides in a host. The method comprises  
CC transforming a host, suitable for fermentation, with genes encoding  
CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.  
CC The specification also describes hosts transformed with a gene active  
CC in glycogen synthesis and at least one non-starch branching gene,  
CC involved in production of amylopectin or amylose in its original host.











```
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 15.3%; Score 298; DB 21; Length 2115;

Best Local Similarity 53.7%; Pred. No. 1.3e-34;
Matches 820; Conservative 0; Mismatches 610; Indels 96; Gaps 6;
QY 20 TTGCTGTGAGTGGCCCTTGGTCCAAAGACGGCGGCTGGGGATGACTGGTGCC 79
Db 415 TTGGAGCTGAAGTTGGTCCATGAGTAATAACTGGTGGTCTTGGTGATCTTCGGTGTGC 474
QY 80 TGCCTATTGAGCTGTCAAGCGCGCCACCGCGTCATGACCAATGCCCTCGCTACGACC 139
Db 475 TACCTCCAGCTCTTGCCTAGAGCCACCGGTGATGACAAATTGTCTCGGTATGACC 534
QY 140 AGTAGCTGACGCTTGGGACACCTCGGTGGTCTGGGACATC-----ATGGCGC 187
Db 535 AATAAAGATGCTTGGGACACCTTGTGTGTGGTTCAGATCAAAAGTTGGGATAAAGTTG 594
QY 188 AGAAGTCCGCTACTTCCACTCCATCAAGAAGGCGTCACCGGTGTGGATTGACCAACC 247
Db 595 AGAATGTCGTTCTTCCATTGCTACAAACGAGGATGATGCTCTTTGTGACCATC 654
QY 248 CCTGTTCTCGGCAAGCTCTGGGGCAAGCCGGTCCAAAGCTGTACGGCCCGGCTCG 307
Db 655 CAATCTTTCTTGTAAAGTTGTGGGCAAAACAGGATCCAAAATCTATGTTCTATAACTG 714
QY 308 GCGCTGACTACTGGACAACCAAGCGCTTGGCCCTCTTCTGCAAGCCCGCTATTGAGS 367
Db 715 GAGTAGACTACAATGACAACCAACTCCGGTTGAGTTTCTGTGTGAGGCTGCTCTGAGG 774
QY 368 CTGCGCGGCTGCTGCGCTTGGCCCGCGGAGGACTG----- 405
Db 775 CACCACAGTCTGAACCTGAACAGCAGCAAGTACTTCTTGGACCATATGTTGAGATG 834
QY 406 --GTCTTGGTGGCAACGACTGGGCACTCCGCCCTGGTGGCCGCTGCTGTGAAGGACGAGT 463
Db 835 TAGTCTTTGTTGCCAATGACTGGCACACTGCTCTACTTCTTACCTTACCTCAATCTATGT 894
QY 464 ACCAGCCCAAGGCCAGTTTCAACAGGCGCAAGTGGTGTGCTGTATCCACACATCGCT 523
Db 895 ATCAATCCCGGAGTCTACATGAATGCAAGGTGGTCTTCTGCAATTCACACATAGCT 954
QY 524 TCCAGGCGCGCATGTGGGAGGAGGCTTTCAAAGGACACGAAGCTGCCCGAGCGCCCTTGG 583
Db 955 ACCAGGGAAGATT-----TGCTTTG 975
QY 584 ACAAGCTGGCTTCTCGGACGCTATGCCAAGTTTACACTGAGCCACCCCATGGAGG 643
Db 976 ATGACTATTCCCTTCTCAACTTGCCTATGAGTTTAAAGTTCTTTCGACTTTCGAGC 1035
QY 644 AGGAGGAGAAGCCCGCTGACGGGAAAGACCTACAAGAAGATCAACTGGTGAAGGCTG 703
Db 1036 GGATGAAAGCC-----AGTAAAGGCGGAAATTAACCTGGATGAAGGCTG 1083
QY 704 GCATTATCGCGCGGACAAAGCTGTGACTGTGTCCCGCAACTACGCGACCGAGATCGCTG 763
Db 1084 CAATTCTGGAAGCTCACCGTGTCTTAAAGTTAGTCCATCTACCTCAAGAATCACT 1143
QY 764 CCGATGCCCGCGGGGTGTGGAGCTGGACACCGTATCCGCCCGGAGGCAATGAGGCA 823
Db 1144 CTGGAGTTGATAGAGCGGTGAATGCAATAATATCTTGAATGAAACAGTTTCCCGAA 1203
QY 824 TTGTGAAGGCGATGACATTCAGGAGTGGAAACCCCAAGACCGACAAGTTCTCTGTGCGC 883
Db 1204 TTATTATGGAATGATGTTCAAGATGGAACCCGCTTACTGACAGTACATCATCATCA 1263
QY 884 CTTACGACCAAGACAGCTCTACGCGGCAAGGCGCGGCAAGGAGGCCCTGACGAGCGG 943
Db 1264 AATACGATATTACCACCTGTTACAGATGCTAAACCATTCATCAAGAAGCACTTCAGGCTG 1323
QY 944 AGCTGGGCTCCTGTGGACCCCGCCCGCTTGTGCTTGTGCTTGTATCGGCGGCTGGAGG 1003
Db 1324 CTGTTGAGCTTCCCGTGGACAGGATGTCCCGGTTATCGGTTTTCATAGAGGATGGAGG 1383
QY 1004 AGCAGAAGGCTGTGACATCATCTCTGGCGGCGCTGCCCAAGATCTCGCCACCCCAAGG 1063
```



QY 809 AGGGCATTGAGGCGATTTGTAACGGCATGGACATTTGAGGAGTGGAAACCCCAAGACCGACA 868  
 Db 1146 AAACCATCTGTGATAGTAATAGTGGATGGACACACCGAGTGAATCCATTAAACAGACA 1205  
 QY 869 AGTTCTGTCTGCGCCCTACGACCAAGACAGCGTCTACGCGGCAAGCGCGCCCAAGG 928  
 Db 1206 AATATATTCTCAAACTACGATGCAACAACTGTATTGGATGCAAAACCTCTCTGTAAAG 1265  
 QY 929 AGGCCCTGAGGCGGAGTGGGCTGCTGTGGACCCCAAGCGCCCTGTTGCGCTTCA 988  
 Db 1266 AAGCTTTGCAAGCTGAGTGGCTGCTGTTTAAACAAAACAAAGCTGTTTGGCCTTG 1325  
 QY 989 TGGCGCCCTGGAGGAGAGAGGTTGGACATATCTTGGCCGCTGCGCCCAAGATCC 1048  
 Db 1326 TTGGAAGACTAGATGAGCAAGAGGCTCAGACATTTAGTGCAGCAATTTCCAGAACTTC 1385  
 QY 1049 TGGCCACCCCAAGTGCAGATCGCCATCTGTTGGTACCGGCAAGCGCGCTACGAGAAGC 1108  
 Db 1386 T---TTGTCAGATGTTCAAGTGATAGTCTGGCACTGGCAAGAAAGTTGGAGAGTG 1442  
 QY 1109 TGGTGAACGCCAPCGGCACCAAGTACAAGGGCGCGCCAGAGGCGGTGGTCAAGTTCTCGG 1168  
 Db 1443 AACTTACATTTAGTGAAGAAATGTTTCCAGACAAATTCAGAGCACATCTCAAAATTCACG 1502  
 QY 1169 CGCCCTGGGCGACATGTCACCGCGCGCGCGGCGGCTATGCTGTTGGCTCGCGCTTCG 1228  
 Db 1503 TTCTTTAGCTATGCAATCATGCGAGAGCTGATATCTCTTGTATTTCCAAAGCAGATTGG 1562  
 QY 1229 AGCCCTGCGGCGTATCCAGCTGCACGCCATGCATACGCTACCGTCCCGTGGTAGCCCT 1288  
 Db 1563 AACCCCTGTGGCCCTCATTCAGCTTCAGGCCAAGCATGATGAACTCTCCCTATGTGTAGCA 1622  
 QY 1289 CCACCGCGCGCTGTCGACACCGTCAAGGAGGCGCTCAGCGGCTTTCACATGGGCGCC- 1347  
 Db 1623 CCACCTGTGACATGTTGACATGTCACAAAGGCTTCACGCTTCCATAGGGCCCT 1682  
 QY 1348 -----CTGAACCCCGCAAGCTGGAGAGCTGACGCGCGAGCGCTGGCGCGCACCGTGC 1402  
 Db 1683 TCAGTGTGGAGTGTGATGCGGTAGACAAAGCTGATGTACAAAAGATTTGCGAAACACGA 1742  
 QY 1403 GCGTGTGCACGAGGTGTTTGGCGGCGCGCTACCCCGAGATGGTGGCCAACTGCAATCA 1462  
 Db 1743 AAGGGCCCTCAAAGTCTATGGAACACCTGCTTTTGTGGAGATGATCAAGAACTGCAATGA 1802  
 QY 1463 GCCAGGACCTGCTCTGTTGTCAGGCGCGCGCCGCAAGTGGGA 1502  
 Db 1803 ACCAAGATCTCTCATGGAAGGACCTGCAAGAAAGTGGGA 1842

## RESULT 8

AAC86435

ID AAC86435 standard; cdna; 2807 BP.

AC AAC86435;

XX 01-MAR-2001 (first entry)

XX Wheat starch synthase clone wssIID cdna.

XX Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;

XX food product; adhesive; ss.

XX Triticum aestivum.

XX WO200066745-A1.

XX 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU000385.

XX 29-APR-1999; 99AU-00000052.

XX (CSIR ) COMMONWEALTH SCI &amp; IND RES ORG.

PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 XX Morell M, Li Z, Rahman S, Appels R;  
 XX WPI; 2000-647602/62.  
 DR Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
 XX WST-II, useful in modifying plant starch content and/or composition -  
 PT disclosure; Fig 2; 21pp; English.  
 PS The present invention relates to novel protein and coding sequences from  
 XX wheat. The proteins are wheat starch synthases, designated SSII and  
 CC SSIII. These can be used in the modification of plant starch content or  
 CC composition, and to screen plants to identify mutations which affect  
 CC starch content and composition. The starch can then be used in food  
 CC products, such as flour, and in films, coatings, adhesives, building  
 CC materials and packaging materials.  
 XX Sequence 2807 BP; 549 A; 844 C; 907 G; 507 T; 0 other;  
 SQ Query Match 11.3%; Score 221; DB 21; Length 2807;  
 Best Local Similarity 56.2%; Pred No. 1.4e-23;  
 Matches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;  
 QY 680 AGAAGATCAACTGCTGGAAGGTTGGCAATATCGCCGCCGAGCAAGCTGGTGTGCG 739  
 Db 1620 AGCAGCGCACTACTTCTGCGCGCGCTGAAGATGCGGACGAGTTGTCTGTGAGCC 1679  
 QY 740 CCAACTAGCGAGGAGATCGTCCGATGCCGCGCGGTGGAGCTGGACACCTCA 799  
 Db 1680 CCGGTTACCTGTGGGAGCTCAAGACGTTGGAGGCGCTGGGGCTTTCAGCATATAC 1739  
 QY 800 TCC---CGCGCAAGCGCATTTGAGGCGATTTGAAACGCGATGACATTTGAGGATGGAAC 856  
 Db 1740 GGCAGACGACTGGAAGACCGCGGCAATGTCACGCGATGCAACATGAGTGGARCC 1799  
 QY 857 CCAAGACGCAAGTTCCTGCTGCGCCCTACGACACCAAGACAGCTGT----- 904  
 Db 1800 CCGAGTTGAGCTCCACTCAAGTCGAGCGCTACACCACTTCTCCCTGGGAGCGTGG 1859  
 QY 905 ACGCCGCGAGGCGCGCGCAAGGAGCGCTGCGAGCGCGGCTGGCGCTGCTGGAC 964  
 Db 1860 ACTCCGCAAGCGCAGTGTGAAGGAGCGCTGCGAGCGCGCTGGCGCTGGCGCTGGCG 1919  
 QY 965 CCACCGCGCGCTGCTGCGCGCTTCATCGCGCGCTGCGAGGACGAGGCTGGACATCA 1024  
 Db 1920 CCGAGTGGCGCTGCTGCGCTTCATCGCGCGCTGCGAGGCGCGCTGGAGATCA 1979  
 QY 1025 TCCTGGCGCGCTGCGCAAGATCTGCGCACCGCCCAAGGTCAGATCGCCATCTGGGTA 1084  
 Db 1980 TCGGGAGCGCATGCCCTGGATCGT---AGCCAGGAGCTGAGCTGATGCTGGGCA 2036  
 QY 1085 CCGCAAGGCGCGCTACGAGAGGCTGTGAACGCCATCGGCACCAAGTACAGGCGCG 1144  
 Db 2037 CCGCGCGCGAGCTGAGAGAGATGCTGGGCGCTTCGAGCGGAGCACCCACAGG 2096  
 QY 1145 CCAAGGCGGTGCTCAAGTTCCTGCGCGCGCTTGGCGCGCATCTCACCGCGCGCGGCT 1204  
 Db 2097 TGGCGGGTGGTGGGTTCTCCGCTGCGCGCTGGCGCGCGGATCACGCGCGCGCGG 2156  
 QY 1205 TCATGCTGGTGGCTCGCGCTTCGAGCGCTGCGCGCTGATCCAGCTGACGCCATGCACT 1264  
 Db 2157 CGCTCTCATGCCCTCGCGCTTCGAGCGCTGCGGCTTGAACACGCTTTACCCATGGCT 2216  
 QY 1265 ACGTACCGTGGCTGGTATGCTTCCACCGCGCGCTGCTGCGACACCGCTCAAGGAGGCG 1324  
 Db 2217 ACGCACCGTGGCTGGTATGCTGCGCGCGCTGCGCGGGTGGAGGAGCACCGCTGGCG 2276  
 QY 1325 TCACCGGCTTCCACATGGCGCGCTGAAACCCCGCAAGCTGGAGGCTGACGCGGACG 1384  
 Db 2277 ACCCTTCAACCACTCGCGCGCTGCGGTGAGGTTGCGACCGCGGAGGCGCAAGCTGA 2336

QY 1385 CCCTGGCCGCCACCGTGGCGGTCCAGAGGAGTGTGTTGGGGCGGCGGTACCCCGAGA 1444  
 Db 2337 TCGAGCGGCTCGGCACTGCTCGGCACCTACCGGACTACAAGGAGAGCTGGAGGGCC 2396  
 QY 1445 TGTGGGCAACTGCATCAGCAGGACCTGCTGTTCCAAAGCCGCCAGAGTGGGAGG 1504  
 Db 2397 TCCAGGAGCGGCGATGTCGAGAGCTTACGTGGGAGCAATGCGGCAAGCTCTACGAGG 2456  
 QY 1505 GCTGCTGTGGAGGAGGTGTGTAC 1527  
 Db 2457 ACCTCTCTCTCAAGGCCAAGTAC 2479

RESULT 9  
 AAC86411  
 ID AAC86411 standard; cDNA; 2842 BP.  
 AC AAC86411;  
 DT 01-MAR-2001 (first entry)  
 XX Wheat starch synthase II coding sequence SEQ ID NO: 3.  
 DE Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;  
 KW food product; adhesive; ss.  
 XX Triticum aestivum.  
 XX WO200066745-A1.  
 PN 09-NOV-2000.  
 XX 28-APR-2000; 2000WO-AU00385.  
 XX 29-APR-1999; 99AU-0000052.  
 XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PA (GOOD-) GOODMAN FIELDER LTD.  
 PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
 XX Morell M, Li Z, Rahman S, Appels R;  
 PI WPI; 2000-647602/62.  
 DR P-PSDB; AAB37567.  
 XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
 PT WST-II, useful in modifying plant starch content and/or composition -  
 XX Claim 1; Page 155-158; 21lpp; English.  
 XX The present invention relates to novel protein and coding sequences from  
 CC wheat. The proteins are wheat starch synthases, designated SSII and  
 CC SSIII. These can be used in the modification of plant starch content or  
 CC composition, and to screen plants to identify mutations which affect  
 CC starch content and composition. The starch can then be used in food  
 CC products, such as flour, and in films, coatings, adhesives, building  
 CC materials and packaging materials.  
 XX Sequence 2842 BP; 561 A; 851 C; 919 G; 511 T; 0 other;  
 SQ Query Match 11.3%; Score 221; DB 21; Length 2842;  
 Best Local Similarity 56.2%; Pred.No.1.4e-23;  
 Matches 485; Conservative 0; Mismatches 360; Indels 18; Gaps 3;

QY 680 AGAAGATCAACTGGCTGAAGGGTGGCATATCGCCCGCAGAGCTGGTGTGTCG 739  
 Db 1620 AGCACCAACTACTTCGCGCGCGGCTAGATGCGGACCAAGTGTGCTGGTGAGCC 1679  
 QY 740 CCAACTACGCGCGAGATGCTGCCGATGCCCGCGGCTGTGGAGCTGACACCGTCA 799  
 Db 1680 CCGGGTACCTGTGGGAGCTCAAGACGCTGGAGGCGGCTTCACGACATCATC 1739

QY 800 TCC---GCGCCAAGGGCAITGAGGSCATTGTCAACGGCATGGACATTGAGGAGTGAACC 856  
 Db 1740 GGCAGAAGCACTGGAAGACCCGCGCATCGTCAAGGGCATGACACATGAGTGAACC 1799  
 QY 857 CCAAGACCGACAAGTTCCTGTCTCGGCCCTAGACCAAGAACAGCGTCT- 904  
 Db 1800 CCGAGGTGGAGTCCACCTCAAGTCGACGGCTACACCACTTCTCCCTGGGGAGCGCTGG 1859  
 QY 905 ACGCGCGCAAGCGCCGCCAAGAGAGCCCTGACGGCGAGCTGGGCTCTGCTGTGGACC 964  
 Db 1860 ACTCGGCAAGCGCAGTGCAGAGAGCCCTGACGGCGAGCTGGGCTCTGAGTTCGCG 1919  
 QY 965 CCACGCGCCCTCTGTTCGCTTCATCGGCCCTCGAGAGAGCAGAAGGCTTGGACATCA 1024  
 Db 1920 CCGACGTGCGCTGCTCGGCTTCATCGGCCCTCGAGCGGCGCAGAAGGCGTGGAGATCA 1979  
 QY 1025 TCCTGGCGGCGCTGCGCCCAAGATCTTGGCCACCCCGAAGTGCATGCCATCTCTGGGTA 1084  
 Db 1980 TCGCGGAGCGCAATGCCCTGGATCGT---AGCCAGGAGCTGACGTGCTATGCTGGCA 2036  
 QY 1085 CCGCAAGGCGCGCTACGAGAAGCTGGTGAACGCCATCGGCACCAAGTACAAAGGCGCG 1144  
 Db 2037 CCGCGCGCCAGCACTGGAGAGCATGCTCGGCACCTTGACCGGGAGCACACGACAAAG 2096  
 QY 1145 CCAAGGGCGTGGTCAAGTTCCTCGCGCCCTCGCGCACATGCTCACCGCGCGCGCGACT 1204  
 Db 2097 TGC CGGGTGGTGGGTTCCTCGCTGCGCTGCGGCACCGGATCACGCGGCGCGCGAGC 2156  
 QY 1205 TCATGCTGTGCTCTCGGCTTCGAGCCCTGCGGCGCTGATCCAGCTGCAGCGCATGCACT 1264  
 Db 2157 CGTCTCTATGCGCTCCCGGTTGAGCGCTGCGGCTTGAACCAAGCTTTAGCCCATGCGCT 2216  
 QY 1265 ACGGTACCGTGGCGGTGTAGCTCCACCGCGCGCTGGTGCACACCGTCAAGGAGGCG 1324  
 Db 2217 ACGGCACTGCGCTGCTGCACCGCTGCGGCGGTGAGGACACCGTGCCTGCGCGCTG 2276  
 QY 1325 TCACCGCTTCCACATGGGCGCTGAACCCCGACAAAGCTGGACGAGCTGACCGCGAGC 1384  
 Db 2277 ACCCTTCAACCACTCGCGCTCGGTGGAGTTCGACCGCGCGGCGGCGCACAAGCTGA 2336  
 QY 1385 CCCTGGCGGCCACCGTGGCGGTGCCAGGAGGTGTTTGGGGCGGCGCTACCCCGAGA 1444  
 Db 2337 TCGAGCGCTCGGGCACTGCTCGCACCTACCGGGACTACAAGGAGAGCTGGAGGGGCC 2396  
 QY 1445 TGGTGGCAACTGCATCAGCAGGACCTGCTCTGCTCCAAAGCCCGCCAGAAGTGGGAGG 1504  
 Db 2397 TCCAGGAGCGGCAATGTCGAGAGCTTACGTGGGAGCATGCCCGCAAGCTCTACGAGG 2456  
 QY 1505 GCTGCTGTGGAGGAGGTGTGTAC 1527  
 Db 2457 ACCTCTCTCTCAAGGCCAAGTAC 2479

RESULT 10  
 AAV01528  
 ID AAV01528 standard; cDNA to mRNA; 2826 BP.  
 XX AAV01528;  
 XX 21-MAY-1998 (first entry)  
 DE Wheat granule-bound starch synthase cDNA clone pTASs1.  
 KW Starch synthase; wheat; transgenic plant; ss.  
 OS Triticum aestivum L. cv. Florida.  
 XX Key Location/Qualifiers  
 FH 162..2561  
 FT CDS /\*tag= a  
 XX WO9745545-A1.













XX 28-APR-2000; 2000WO-AU00385.  
XX 29-APR-1999; 99AU-0000052.  
XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
XX (GOOD-) GOODMAN FIELDER LTD.  
XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
XX Morell M, Li Z, Rahman S, Appels R;  
XX WPI; 2000-647602/62.  
XX  
XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
XX WST-II, useful in modifying plant starch content and/or composition -  
XX Claim 1; Page 195-200; 21lpp; English.  
XX  
XX The present invention relates to novel protein and coding sequences from  
XX wheat. The proteins are wheat starch synthases, designated SSII and  
XX SSIII. These can be used in the modification of plant starch content or  
XX composition, and to screen plants to identify mutations which affect  
XX starch content and composition. The starch can then be used in food  
XX products, such as flour, and in films, coatings, adhesives, building  
XX materials and packaging materials.  
XX  
XX Sequence 9024 BP; 2228 A; 2386 C; 2216 G; 2193 T; 1 other;  
Query Match 10.9%; Score 212.4; DB 21; Length 9024;  
Best Local Similarity 55.7%; Pred. No. 2.2e-22;  
Matches 477; Conservative 0; Mismatches 361; Indels 18; Gaps 3;  
QY 687 CAACCTGGCTGAAGGTGGCATATCCCGCCGACACAGCTGGTACTGTGCGCCCAACTA 746  
DB 7392 CAACACTACTTGGCGCGGCTGAAGATGGCGGACCACTGTGCTGGTGGAGCCCGGGTA 7451  
QY 747 CGCGACCGAGATCGCTGCCGATGCCCGCGGCTGGAGCTGGACACGCTCATCC- -G 803  
DB 7452 CCGTGTGGAGCTCAAGACGCTGGAGGCGGCTGGGGCTTCACGACATATAGGCGAGAA 7511  
QY 804 CGCCAGGCGCATGGAGGCAATTGTGAACGGCATGGAGATTGAGAGTGGAAACCCCAAGAC 863  
DB 7512 CGACTGGAAGACCGCGGCGCATGCTCAACGGCATCGACACATGGATGGAAACCCCGAGGI 7571  
QY 864 CGACAAGTTCTGTGCGCCCTTACGACACGACAGCGTCT- - - - -ACGCCGG 911  
DB 7572 GGACGCCACCTCAAGTCGGAGGGCTACACCACTTCTCCCTGAGGACGCTGGACTCCGG 7631  
QY 912 CAAGCGCGCGCCAGGAGCGCTCGAGGCGGAGCTGGGCGCTGCCCTGGAGCCCAACGC 971  
DB 7632 CAAGCGCGAGTCAAGGAGGCGCTCGAGCGGAGCTGGGCGCTGCAGTCCGCGCGACGT 7691  
QY 972 CCCCCTGTGCGCTTCATCGGCGCGCTGGAGGAGCAGAGGCTGTGACATCATCCTGGC 1031  
DB 7692 GCGCTGTGCGCTTCATCGGCGCGCTGGAGCGGAGAGGCGTGGAGATCATCGCGA 7751  
QY 1032 CGCCCTGCGCAAGATCCTGGCCACCCCAAGTGCAGATCGCCATCCTGGGTACCGGAA 1091  
DB 7752 CGCCATGCGCTGGATCGTG- - -AGCAGGACGTGAGCTGGTGTGATGCTGGGACCGGGG 7808  
QY 1092 GCGCCCTACGAGAGCTGGTGAACGCCATCGGACCAAGTACAGGGCGCGCCAAAGG 1151  
DB 7809 CCACGACCTGGAGAGCATGCTCGGCACTTCAGGCGGAGACACGACAGGTGCGCGG 7868  
QY 1152 CGTGTCAAGTTCTCGGCGCGCTGGCGACATGCTCAACGCGCGCGGCGGACTTCATGCT 1211  
DB 7869 GTGGGTGGGGTCTCCGTGCGCTGGCGACCGGATCACGGCGGGCGGCGGCGCTCCT 7928  
QY 1212 GGTGCGCTCGGCTTCGAGCCCTGGGCTGATCCAGCTGACGCGCATGCACTACGGTAC 1271  
DB 7929 CATGCCCTCCCGGTTTCGAGCGGCTGGGCTGAACCAAGCTCTACGCCATGGCTACGGAC 7988  
QY 1272 CGTGCCCGTGTAGCTCCACCGGCGGCGCTGGTGCACACCGCTCAAGGAGGGCGTCAACGG 1331

DB 7989 CGTCCCGCTCGTGACGCGCTTGGCGCTCAGGAGACACCGTGGCGCGTTTCGACCCCT 8048  
QY 1332 CTTCCACATGGCGCGCTTGAACCCCGACAACTGGACAGAGCTGACGCCGAGCGCTGGC 1391  
DB 8049 CAACCACTCCGGGCTCGGGTGGAGCTTCGACCGCGCGGAGCGGACAAAGCTGATCGAGG 8108  
QY 1392 CGCCACCGTGGCGCGCTGCCAGCGAGGTCTTTTGGGGGCGCGCTACCCCGAGATGGTGC 1451  
DB 8109 GCTCGGCACTGCTCCGCACTACCGAGACTTCAAGGAGAGCTGGAGGCGCTCCAGGA 8168  
QY 1452 CAACCTCATCAGCCAGCAAGCTGCTCTGTCCAAAGCCCGCCAGAGTGGGAGGCGCTGCT 1511  
DB 8169 GCGCGCATGCTCGAGGACTTCACTGGGAGCACGCCCAAGCTCTACGAGGAGCTCT 8228  
QY 1512 GGAGGAGGTGGTGTAC 1527  
DB 8229 CGTCAAGGCCAAGTAC 8244

Search completed: June 3, 2003, 10:08:08  
Job time : 344.834 secs

***This Page Blank (uspto)***

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:33:29 ; Search time 1905.46 Seconds  
(without alignments)  
16599.542 Million cell updates/sec

Title: US-09-980-771a-4

Perfect score: 1953

Sequence: 1 gcgcggacatcgtatggt.....ccgccaccgcgaagtcgcgc 1953

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508.6	26.0	786	12	BF864001
2	476.6	24.4	502	14	BQ814687
3	469.8	24.1	473	10	AV640741
4	457.8	23.4	461	10	AV629488
5	456	23.3	666	13	BM003295
6	454.6	23.3	2147	11	AY109531

7	444.4	22.8	458	10	AV622787
8	400.8	20.5	436	10	AV629117
9	395.6	20.3	511	10	AV394077
10	391.4	20.0	393	10	AV631115
11	385.2	18.7	475	10	AM757933
12	352.6	18.1	367	13	BI727902
13	351.2	18.0	799	12	BF863935
14	337.8	17.3	418	10	AV644765
15	295.8	15.1	506	10	AV643281
16	278.4	14.3	702	14	BQ246353
17	272.8	14.0	560	10	AV628313
18	251.8	12.9	552	10	BE024926
19	250.2	12.8	517	10	AV641583
20	245.6	12.6	726	14	BQ04991
21	244.2	12.5	532	10	AV641989
22	242	12.4	585	14	BQ246366
23	235.8	12.1	513	10	AV631004
24	235.8	12.1	552	10	AV640546
25	234.8	12.0	524	10	AV641724
26	234.2	12.0	500	10	AV642869
27	231.8	11.9	510	10	AV395307
28	230.8	11.8	761	12	RG366176
29	223.8	11.5	540	10	AV642722
30	222.2	11.4	1545	11	AY112448
31	216.6	11.1	773	12	BQ351175
32	212	10.9	572	14	BQ245812
33	211.8	10.8	486	10	AV642834
34	211.8	10.8	504	10	AV629198
35	210.8	10.8	502	10	AV644517
36	210.2	10.8	511	10	BE423250
37	209.2	10.7	696	12	BG351920
38	208.8	10.7	495	10	AV642565
39	208.2	10.7	543	14	BQ245428
40	207	10.6	641	14	BQ247154
41	206.8	10.6	522	10	BE423625
42	206.4	10.6	729	12	BG599615
43	200.8	10.3	493	10	AV631838
44	200.2	10.3	2498	11	AY109714
45	200	10.2	631	13	BI531619

#### ALIGNMENTS

RESULT 1  
BF864001 786 bp mRNA linear EST 19-JAN-2001  
LOCUS 963048D01.yl C. reinhardtii CC-1690, Stress condition I, normalized  
DEFINITION , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
ACCESSION BF864001  
VERSION BF864001.1 GI:12254145  
KEYWORDS EST.  
SOURCE Chlamydomonas reinhardtii.  
ORGANISM Chlamydomonas reinhardtii.  
REFERENCE 1 (bases 1 to 786)  
AUTHORS Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.  
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3  
JOURNAL Unpublished (2000)  
COMMENT Contact: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu  
LOCATION/Qualifiers  
1. .786  
/organism="Chlamydomonas reinhardtii"

```
/strain="CC-1690 wild type mt+ 21gr"  
/db_xref="taxon:3055"
```

```
/clone_lib="C. reinhardtii CC-1690, Stress condition I,  
normalized, Lambda Zap II"  
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:  
XhoI; This library, constructed by John Davies and Jeffrey  
McDermott, combines cDNAs from CC-1690 cells grown to  
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,  
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr  
, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was  
purified from each sample, pooled and cDNA synthesized.  
The cDNA was directionally cloned into lambda Zap II  
(Stratagene) in the EcoRI (5') and XhoI (3') sites.  
pBluescript II SK- plasmids were excised from the lambda  
ZAP clones by superinfection with ExAssist (Stratagene)  
phage. The library was normalized using method 4 described  
in Bonaldo et al (1996) Genome Research 6: 791-806."
```

```
BASE COUNT 157 a 282 c 244 g 102 t 1 others
```

## ORIGIN

```
Query Match 26.0%; Score 508.6; DB 12; Length 786;  
Best Local Similarity 95.3%; Pred. No. 2e-77; Mismatches 24; Indels 3; Gaps 2;  
Matches 546; Conservative 0;
```

QY 1383 CGCCCTGGCGCCACCGTGGCGGTGCGAGCGAGGTGTTTGGGGGGCGGCTACCCCGA 1442  
Db 95 CCCCCGGGTGCAAGATTGGGCACGAGGCGAGGTGTTTGGGGGGCGGCTACCCCGA 154  
QY 1443 GATGTGGCAACTGATCAGCAGGACCTGCTGCTCAAGCCCGCCAGAGTGGGA 1502  
Db 155 GATGGTGGCAACTGATCAGCAGGACCTGCTGCTCAAGCCCGCCAGAGTGGGA 214  
QY 1503 GGCCCTGCTGGAGAGTGTGACGCAAGGGCGGTGGCCACCGCCCAAGAGGAGGA 1562  
Db 215 GGGCTGCTGGAGAGTGTGACGCAAGGGCGGTGGCCACCGCCCAAGAGGAGGA 274  
QY 1563 GATCAAGTGGCCGTTGCCGAGAGATGCCCGGCGAGCTGCCCGCGGTGCTCTACGCCCC 1622  
Db 275 GATCAAGTGGCCGTTGCCGAGAGATGCCCGGCGAGCTGCCCGCGGTGCTCTACGCCCC 334  
QY 1623 CACACCTCAAGCCGTCGCGCTCCGTGGAGGCAAGCGGCGCGCCGCCCAAGGT 1682  
Db 335 CACACCTCAAGCCGTCGCGCTCCGTGGAGGCAAGCGGCGCGCCGCCCAAGGT 394  
QY 1683 CGGCACCAACCGCCCGCCATGGCGGTGGCGGCGGACACCCCTCGGGCCCTCGCC 1742  
Db 395 CGGCACCAACCGCCCGCCATGGCGGTGGCGGCGGACACCCCTCGGGCCCTCGCC 454  
QY 1743 CGCGCGCGCCACCGCCCAAGGTACCACTACAGCCCGCGCGCTGCCCGCCACCGCAAGCC 1802  
Db 455 CGCGCGCGCCACCGCCCAAGGTACCACTACAGCCCGCGCGCTGCCCGCCACCGCAAGCC 514  
QY 1803 CAAGACCGCTGGCTCAAGCTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGG 1862  
Db 515 CAGACCGCTGGCTCAAGCTGGCGGTGAGGCTCCACACCTCGACCTCGGAGAACGG 574  
QY 1863 CGCTGCTCCACAGGCAACGGCAACGGTGGCTCGGGCTCCCAAGACCTCGGCTGCCAAGCC 1922  
Db 575 CGCTGCTCCCAAGCGC-ACGGCAACGGTGGCTCGGGCTCCCAAGACCTCGGCTGCCAAGCC 633  
QY 1923 CTGGGT--CTCCGCGCCACCGCAAGTCCGCC 1953  
Db 634 CTGGGTCTTCGGGGGCGCCACCGGAAGTTCCGCC 666

## RESULT 2

```
BQ814687  
LOCUS BQ814687 502 bp mRNA linear EST 01-AUG-2002  
DEFINITION 1030045B10.y1 C. reinhardtii CC-1690, Deflagellation (normalized),  
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
ACCESSION BQ814687  
VERSION BQ814687.1 GI:22062228  
KEYWORDS EST.
```

SOURCE  
ORGANISM

```
Chlamydomonas reinhardtii.  
Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadaeae; Chlamydomonas.
```

## REFERENCE

1 (bases 1 to 502)  
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.  
Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
Unicellular System for Analyzing Gene Function and Regulation in  
Vascular Plants. Project: 1030

## JOURNAL

Unpublished (2002)

## COMMENT

Contact: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu.

## FEATURES

source

```
1..502  
/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db_xref="taxon:3055"  
/clone_lib="C. reinhardtii CC-1690, Deflagellation  
(normalized), Lambda Zap II"  
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:  
XhoI; Deflagellation library, constructed by John Davies  
and Jeffrey McDermott, combines cDNAs from CC-1690 cells  
which had been re-synthesizing flagella for 15, 30 and 60  
min after being deflagellated by pH shock. PolyA mRNA was  
purified from each sample, pooled and cDNA synthesized.  
The cDNA was directionally cloned into lambda Zap II  
(Stratagene) in the EcoRI (5') and XhoI (3') sites.  
pBluescript II SK- plasmids were excised from the lambda  
ZAP clones by superinfection with ExAssist (Stratagene)  
phage. The library was normalized using method 4 described  
in Bonaldo et al., (1996) Genome Research 6: 791-806."
```

```
BASE COUNT 79 a 200 c 164 g 59 t
```

## ORIGIN

```
Query Match 24.4%; Score 476.6; DB 14; Length 502;  
Best Local Similarity 98.2%; Pred. No. 5.6e-72;  
Matches 482; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

QY 1353 CCCCAGCAAGCTGACGAGGCTGACCGCGAGCGCCTGCGCCCGCCAGCTGCGCCGCGCAG 1412  
Db 12 CCGTGTGCGGTGGAGGCTGACCGCGAGCGCCTGCGCCCGCCAGCTGCGCCGCGCAG 71  
QY 1413 CGAGGTGTTTGGCGGCGCGCTACCCCGAGATGTTGGCCCAACTGCATCAGCAGGACCT 1472  
Db 72 CGAGGTGTTTGGCGGCGCGCTACCCCGAGATGTTGGCCCAACTGCATCAGCAGGACCT 131  
QY 1473 GTCCTGTGTCCAAAGCCCGCCAGAGTGGAGGGCTGCTGGAGGAGTGGTGTACGGCAA 1532  
Db 132 GTCCTGTGTCCAAAGCCCGCCAGAGTGGAGGGCTGCTGGAGGAGTGGTGTACGGCAA 191  
QY 1533 GGGCGGCGTGGCCACCGCCCAAGAGGAGGATCAAGTGGCCGCTGCGGAGACATCCC 1592  
Db 192 GGGCGGCGTGGCCACCGCCCAAGAGGAGGATCAAGTGGCCGCTGCGGAGAGATCCC 251  
QY 1593 CGCGACCTGCGCGCGGTGCTCTACGCCCGCCCAACACCTGAAGCCGCTGCTCGGCTCCGT 1652  
Db 252 CGCGACCTGCGCGCGGTGCTCTACGCCCGCCCAACACCTGAAGCCGCTGCTCGGCTCCGT 311  
QY 1653 GGAGGGCAACGGCGCGCGCGCCCAAGTCTGGCACCCACCGCCCGCCCGCATGGGCGCGTG 1712  
Db 312 GGAGGGCAACGGCGCGCGCGCCCAAGTCTGGCACCCACCGCCCGCCCGCATGGGCGCGTG 371  
QY 1713 CGCGCGCACACCGCCCTCGGCGCCCTCGCGCGCGCGCCCGCCCGCCCGCCCGCCCGCA 1772  
Db 372 CGCGCGCACACCGCCCTCGGCGCCCTCGCGCGCGCGCCCGCCCGCCCGCCCGCCCGCA 431  
QY 1773 CAAGCCCGCCCTGCGCGCCACCGCCCAAGACCGCTGCCTCAAGCTGGCGGTGA 1832

```

||||| 432 CAAGCCGCTGTCGCGCCACCGTCAAGCCCAAGACCGCTGGCTCAAGCTGGCCGGTGA 491
QY 1833 GGCCTCCACCA 1843
Db 492 GGCCTCCACCA 502

RESULT 3
AV640741 AV640741 473 bp mRNA linear EST 15-DEC-2000
LOCUS AV640741 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION cDNA clone HCL021f06_r 5', mRNA sequence.
ACCESSION AV640741
VERSION AV640741.1 GI:10784069
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Chlamydomonas reinhardtii.
AUTHORS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
1 (bases 1 to 473) Chlamydomonadaceae; Chlamydomonas.
Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
TITLE DNA Res. 7 (5), 305-307 (2000)
JOURNAL 20539644
MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..473
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone_lib="HCL021f06_r"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT 76 a 162 c 146 g 89 t
ORIGIN

Query Match 24.1%; Score 469.8; DB 10; Length 473;
Best Local Similarity 99.6%; Pred. No. 8.1e-71;
Matches 471; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 TGGTTGCTGCTGAGGTGCGCCCTGGTCCCAAGACGGCGGCGGTGGCGATGTGACTGGTG 76
Db 1 TGGTTGCTGCTGAGTGTGCGCCCTGGTCCCAAGACGGCGGCTGGCGATGTGACTGGTG 60

QY 77 GCGTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACCATGCCCCCTCGCTACG 136
Db 61 GCGTGCTATTGAGCTGGTCAAGCGGGCCACCGCGTCAATGACCATGCCCCCTCGCTACG 120

QY 137 ACCAGTACGCTGAGCGCTGGGACACCTCGGTGTGTGGACATCATGGGCGAGAAGTCC 196
Db 121 ACCAGTACGCTGAGCGCTGGGACACCTCGGTGTGTGGACATCATGGGCGAGAAGTCC 180

QY 197 GCTACTTCCACTCCATCAAGAAGGGGTGCACCGCGTGTGGATGTGACCCCGCTGGTTC 256
Db 181 GCTACTTCCACTCCATCAAGAAGGGGTGCACCGCGTGTGGATGTGACCCCGCTGGTTC 240

QY 257 TGCCCAAGGTCTTGGGCAAGACCGGTCCAAAGCTGTACGGCCCGCTCGCGGCGCTGACT 316
Db 241 TGCCCAAGGTCTTGGGCAAGACCGGTCCAAAGCTGTACGGCCCGCTCGCGGCGCTGACT 300

QY 317 ACCTGACACACCAAGCGGCTTCGCCCTGTGTGCAAGGCCCGCTATTGAGGCTGCCCGG 376
Db 301 ACCTGACACACCAAGCGGCTTCGCCCTGTGTGCAAGGCCCGCTATTGAGGCTGCCCGG 360

```

```

QY 377 TGCTGCCCTTCGGCCCGCGGAGGACTGGTCTTGTTGGCCCAAGGACTGGCACTCGGCC 436
Db 361 TGCTGCCCTTCGGCCCGCGGAGGACTGGTCTTGTTGGCCCAAGGACTGGCACTCGGCC 420

QY 437 TGGTCCCGCTGCTGCTGAAGGACGAGTACCAAGGCCCAAGGCCAGTTTCAACAAG 489
Db 421 TGGTCCCGCTGCTGCTGAAGGACGAGTACCAAGGCCCAAGGCCAGTTTCAACAAG 473

RESULT 4
AV629488 AV629488 461 bp mRNA linear EST 15-DEC-2000
LOCUS AV629488 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas
DEFINITION reinhardtii cDNA clone LCL059f09_r 5', mRNA sequence.
ACCESSION AV629488
VERSION AV629488.1 GI:10792122
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Chlamydomonas reinhardtii.
AUTHORS Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
1 (bases 1 to 461) Chlamydomonadaceae; Chlamydomonas.
Asamizu,E., Miura,K., Kucho,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
Nakamura,Y. and Tabata,S.
Generation of expressed sequence tags from low-CO2 and high-CO2
adapted cells of Chlamydomonas reinhardtii
TITLE DNA Res. 7 (5), 305-307 (2000)
JOURNAL 20539644
MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..461
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone_lib="LCL059f09_r"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a carbon stress acclimatized condition in which carbon
dioxide concentration in the bubbling gas was changed from
5% to 0.04%"
BASE COUNT 73 a 156 c 143 g 89 t
ORIGIN

Query Match 23.4%; Score 457.8; DB 10; Length 461;
Best Local Similarity 99.6%; Pred. No. 9.1e-69;
Matches 459; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 ACATCGGTGATGGTGTGCTGAGTCCGCCCTTGCTCAAGACGGCGGCGCTGGGCCGATG 67
Db 1 ACATCGGTGATGGTGTGCTGAGTCCGCCCTTGCTCAAGACGGCGGCGCTGGGCCGATG 60

QY 68 TGACTGTGGCGCTCCCTTATTTAGCTGCTCAAGCGCGGCCACCGCGTGCATGACCATTCGCC 127
Db 61 TGACTGTGGCGCTCCCTTATTTAGCTGCTCAAGCGCGGCCACCGCGTGCATGACCATTCGCC 120

QY 128 CTCGCTACGACCAAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGGACATCATGGCG 187
Db 121 CTCGCTACGACCAAGTACGCTGACGCTGGGACACCTCGGTGGTGGTGGACATCATGGCG 180

QY 188 AGAAGTCCCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGGTGTGATTTGACCAACC 247
Db 181 AGAAGTCCCGCTACTTCCACTCCATCAAGAAGGGGTGCACCGGTGTGATTTGACCAACC 240

QY 248 CTTGGTTCCTGGGCCAAGGTCTGGGGCAAGACCGGCTCCAAAGCTGTAGGGCCCGGTCGG 307
Db 241 CTTGGTTCCTGGGCCAAGGTCTGGGGCAAGACCGGCTCCAAAGCTGTAGGGCCCGGTCGG 300

```

QY 308 GCGCTGACTACCTGGACCAACCAAGCGCTTGGCCCTGTCTTTCGCAAGGCGGCTATTGAGG 367  
 |||||  
 Db 301 GCGCTGACTACCTGGACCAACCAAGCGCTTGGCCCTGTCTTTCGCAAGGCGGCTATTGAGG 360  
 |||||  
 QY 368 CTGCGCGCGTGTGCGCCCTTGGCGCCCGGAGGACTGCGTCTTCGTGGCCCAAGACTGGC 427  
 |||||  
 Db 361 CTGCGCGCGTGTGCGCCCTTGGCGCCCGGAGGACTGCGTCTTCGTGGCCCAAGACTGGC 420  
 |||||  
 QY 428 ACTCCGCGCTGTGCGCGCGTCTCTCTCAAGGACGAGTACCAG 468  
 |||||  
 Db 421 ACTCCGCGCTGTGCGCGCGTCTCTCTCAAGGACGAGTACCAG 461  
 |||||

RESULT 5  
 LOCUS BM003295 666 bp mRNA linear EST 25-OCT-2001  
 DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.  
 ACCESSION BM003295  
 VERSION BM003295.1 GI:16438075  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii.  
 ORGANISM Chlamydomonas reinhardtii.  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadales; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 666)  
 AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre  
 P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.  
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,  
 Unicellular System for Analyzing Gene Function and Regulation in  
 Vascular Plants. Project: 1031  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Charles Hauser  
 DCMB Box 91000  
 Duke University  
 Durham, NC 27708-1000  
 Tel: 919 613 8159  
 Fax: 919 613 8177  
 Email: chauser@duke.edu.

FEATURES  
 source  
 1..666  
 /organism="Chlamydomonas reinhardtii"  
 /strain="CC-1690 wild type mt+ 21gr"  
 /db\_xref="taxon:3055"  
 /clone\_lib="C. reinhardtii CC-1690, Stress II (normalized  
 ), Lambda Zap II"  
 /note="Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Stress condition II library, constructed by John  
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
 cells grown to mid-log phase in TAP (NH4+ - containing)  
 and shifted to TAP - NO3- (24hrs); H2 production  
 conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant  
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
 PolyA mRNA was purified from each sample, pooled and cDNA  
 synthesized. The cDNA was directionally cloned into lambda  
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')  
 sites. pBluescript II SK- plasmids were excised from the  
 lambda Zap clones by superinfection with ExAssist  
 (Stratagene) phage. The library was normalized using  
 method 4 described in Bonaldo et al., (1996) Genome  
 Research 6: 791-806."  
 BASE COUNT 121 a 244 c 224 g 77 t  
 ORIGIN

Query Match 23.3%; Score 456; DB 13; Length 666;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-68;  
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1498 TGGGAGGCGCTGTGGAGAGGTGGTGTACGGCAAGGGCGGCTGGCCACCGCAAGAAG 1557  
 |||||  
 Db 1 TGGGAGGCGCTGTGGAGAGGTGGTGTACGGCAAGGGCGGCTGGCCACCGCAAGAAG 60  
 |||||

QY 1558 GAGGAGATCAAGTGTGCCGTTCGCGAGAAGATCCCCGGGACCTGCCCGCGGTCTCTAC 1617  
 |||||  
 Db 61 GAGGAGATCAAGTGTGCCGTTCGCGAGAAGATCCCCGGGACCTGCCCGCGGTCTCTAC 120  
 |||||  
 QY 1618 GCCCCCAACACCTTGAAGCCCGTGTCCGCTCCGTGAGGGGACACGGCGCGCGCC 1677  
 |||||  
 Db 121 GCCCCCAACACCTTGAAGCCCGTGTCCGCTCCGTGAGGGGACACGGCGCGCGCC 180  
 |||||  
 QY 1678 AAGTGTGGCACCAACCGCCCGCCATGGCGCGTGGCGCGACACCCCTCTCGGGCCCC 1737  
 |||||  
 Db 181 AAGTGTGGCACCAACCGCCCGCCATGGCGCGTGGCGCGACACCCCTCTCGGGCCCC 240  
 |||||  
 QY 1738 TCGCCCGCCCGCCCGCCCGCCAGGTGACCACTACAAAGCCCGCCCTGCCCGCACCGCC 1797  
 |||||  
 Db 241 TCGCCCGCCCGCCCGCCCGCCAGGTGACCACTACAAAGCCCGCCCTGCCCGCACCGCC 300  
 |||||  
 QY 1798 AAGCCCAAGACCGCTGCGCTCAAGCTGCGCGTGGAGCCCTCCACACCTGACCTCGGAG 1857  
 |||||  
 Db 301 AAGCCCAAGACCGCTGCGCTCAAGCTGCGCGTGGAGCCCTCCACACCTGACCTCGGAG 360  
 |||||  
 QY 1858 AAGCGCGCTGCTCCAAACGGCAACGGCAAGGTGCTTCGGCTCCAAAGACCTCGGCTGCC 1917  
 |||||  
 Db 361 AAGCGCGCTGCTCCAAACGGCAACGGCAAGGTGCTTCGGCTCCAAAGACCTCGGCTGCC 420  
 |||||  
 QY 1918 AAGCCCTGTGCTCCCGCGCCCGCCGCAAGTCCGCC 1953  
 |||||  
 Db 421 AAGCCCTGTGCTCCCGCGCCCGCCGCAAGTCCGCC 456  
 |||||

RESULT 6  
 LOCUS AY109531 2147 bp mRNA linear HTC 25-MAY-2002  
 DEFINITION Zea mays CL1198\_1 mRNA sequence.  
 ACCESSION AY109531  
 VERSION AY109531.1 GI:21213285  
 KEYWORDS HTC.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 2147)  
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 JOURNAL Unpublished (2002)  
 REFERENCE 2 (bases 1 to 2147)  
 AUTHORS Coe, E.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA

FEATURES  
 source  
 1..2147  
 /organism="Zea mays"  
 /db\_xref="MaizeDB:630100"  
 /db\_xref="taxon:4577"  
 /clone="CL1198\_1"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

BASE COUNT 400 a 582 c 604 g 365 t 196 others  
 ORIGIN  
 Query Match 23.3%; Score 454.6; DB 11; Length 2147;  
 Best Local Similarity 58.6%; Pred. No. 3.8e-68;  
 Matches 910; Conservative 0; Mismatches 547; Indels 96; Gaps 6;





Matches 456; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 365 AGGCTGCCCGCTGCTGCCCTTCGGCCGGGAGGAGTCTCTTCGTTGGCCCAAGCACT 424  
 |||||  
 Db 1 AGGCTGCCCGCTGCTGCCCTTCGGCCGGGAGGAGTCTCTTCGTTGGCCCAAGCACT 60  
 |||||

QY 425 GGCACCTCCGCCCTGGTGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 484  
 |||||  
 Db 61 GGCACCTCCGCCCTGGTGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 |||||

QY 485 CCAAGGCCAAGTCTGGTGGTCTATC-CACATATCTGCTCTCTCTCTCTCTCTCTCTCTCT 543  
 |||||  
 Db 121 CCAAGGCCAAGTCTGGTGGTCTATCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 |||||

QY 544 GAGGCTTCAAGGACACGAGTCTGCCAGGCGGCTTTGACAAGCTGGCTTCTCTCTCTCT 603  
 |||||  
 Db 181 GAGGCTTCAAGGACATGAGTCTGCCAGGCGGCTTTGACAAGCTGGCTTCTCTCTCTCT 240  
 |||||

QY 604 GGCTATGCCAAGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAACGCCCGCTG 663  
 |||||  
 Db 241 GGCTATGCCAAGTTTACACTGAGGCCACCCCATGAGGAGGACGAGAACGCCCGCTG 300  
 |||||

QY 664 ACGGGAACACCTACAGAGATCACTGGCTGAAGGGTGGCATTTATCGCGCCGCAAG 723  
 |||||  
 Db 301 ACGGGAACACCTACAGAGATCACTGGCTGAAGGGTGGCATTTATCGCGCCGCAAG 360  
 |||||

QY 724 CTGGTACTGTGTCGCCCACTACGCGACGAGATCGCTGCGATCGCGCGGCTGTG 783  
 |||||  
 Db 361 CTGGTACTGTGTCGCCCACTACGCGACGAGATCGCTGCGATCGCGCGGCTGTG 420  
 |||||

QY 784 GAGCTGGACACCTCATCGCGCCCAAGGGCAATTGAGGG 821  
 |||||  
 Db 421 GAGCTGGACACCTCATCGCGCCCAAGGGCAATTGAGGG 458  
 |||||

RESULT 8  
 AV629117 436 bp mRNA linear EST 15-DEC-2000  
 LOCUS AV629117 Chlamydomonas reinhardtii 5% to 0.04% CO2 Chlamydomonas  
 DEFINITION reinhardtii cDNA clone LCL052d05\_r 5', mRNA sequence.  
 ACCESSION AV629117  
 VERSION AV629117.1 GI:10791751  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii.  
 ORGANISM Chlamydomonas reinhardtii.  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 436)  
 AUTHORS Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,  
 Nakamura, Y., and Tabata, S.  
 TITLE Generation of expressed sequence tags from low-CO2 and high-CO2  
 adapted cells of Chlamydomonas reinhardtii  
 JOURNAL DNA Res. 7 (5), 305-307 (2000)  
 MEDLINE 20539644  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
 source  
 1. .436  
 /organism="Chlamydomonas reinhardtii"  
 /strain="C9"  
 /db\_xref="taxon:3055"  
 /clone="LCL052d05\_r"  
 /clone\_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"  
 /note="Vector: pBluescriptII SK; Site\_1: EcoRI; Site\_2:  
 XhoI. The cDNA library was constructed from cells cultured  
 in a carbon stress acclimatized condition in which carbon  
 dioxide concentration in the bubbling gas was changed from  
 5% to 0.04%"

BASE COUNT 66 a 145 c 138 g 87 t

Query Match 20.5%; Score 400.8; DB 10; Length 436;  
 Best Local Similarity 99.5%; Pred. No. 5e-59;  
 Matches 402; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCCTGGACATCGTGTGCTGCTGAGGTGCGCCCTTGGTCTCCAGAGCGGCGGCTG 60  
 |||||  
 Db 33 GGCCTGGACATCGTGTGCTGCTGAGGTGCGCCCTTGGTCTCCAGAGCGGCGGCTG 92  
 |||||

QY 61 GGCATGTGACTGTGGCTGCTGCTATTGAGCTGTCAAGCGCGCCACCGCTCATGACC 120  
 |||||  
 Db 93 GGCATGTGACTGTGGCTGCTGCTATTGAGCTGTCAAGCGCGCCACCGCTCATGACC 152  
 |||||

QY 121 ATTGCCCTCGCTGCTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 |||||  
 Db 153 ATTGCCCTCGCTGCTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212  
 |||||

QY 181 ATGGGGGAGAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 |||||  
 Db 213 ATGGGGGAGAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 272  
 |||||

QY 241 GACCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 |||||  
 Db 273 GACCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332  
 |||||

QY 301 CGCTCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 |||||  
 Db 333 CGCTCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392  
 |||||

QY 361 ATTGAGGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404  
 |||||  
 Db 393 ATTGAGGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436  
 |||||

RESULT 9  
 AV394077 511 bp mRNA linear EST 23-APR-2002  
 LOCUS AV394077 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii  
 DEFINITION cDNA clone CL18c08\_r 5', mRNA sequence.  
 ACCESSION AV394077  
 VERSION AV394077.1 GI:6548293  
 KEYWORDS EST.  
 SOURCE Chlamydomonas reinhardtii.  
 ORGANISM Chlamydomonas reinhardtii.  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 REFERENCE 1 (bases 1 to 511)  
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H., and Tabata, S.  
 TITLE A large scale structural analysis of cDNAs in a unicellular green  
 alga, Chlamydomonas reinhardtii. I. Generation of 3433  
 non-redundant expressed sequence tags  
 JOURNAL DNA Res. 6 (6), 369-373 (1999)  
 MEDLINE 20152988  
 COMMENT Contact: Yasukazu Nakamura  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
 source  
 1. .511  
 /organism="Chlamydomonas reinhardtii"  
 /strain="C9"  
 /db\_xref="taxon:3055"  
 /clone="CL18c08\_r"  
 /clone\_lib="Chlamydomonas reinhardtii C9"  
 /dev\_stage="photoautotrophic growth"  
 /note="Vector: pBluescriptII SK; Site\_1: EcoRI; Site\_2:  
 XhoI"

BASE COUNT 80 a 170 c 159 g 101 t 1 others

Query Match 20.3%; Score 395.6; DB 10; Length 511;  
 Best Local Similarity 98.8%; Pred. No. 3.9e-58;





## FEATURES

MEDLINE

```
source
1. .506
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="HCL066h09_r"
/clone_lib="Chlamydomonas reinhardtii 5% CO2"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      77 a  170 c  160 g   99 t
ORIGIN
Query Match      15.1%; Score 295.8; DB 10; Length 506;
Best Local Similarity 99.3%; Pred. No. 4.5e-41;
Matches 297; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCGCTGGACATCGTGAATGTTGCTGTAGTGGCGCCCTTGGTCCAAAGAGCGGCGCGCTG 60
Db 208 GCGCTGGACATCGTGAATGTTGCTGTAGTGGCGCCCTTGGTCCAAAGAGCGGCGCTG 267
QY 61 GCGCATGTGACTGTGGCTGCTTATGAGTGGTCAAGCGGCGCCACCGCGTCATGACC 120
Db 268 GCGCATGTGACTGTGGCTGCTTATGAGTGGTCAAGCGGCGCCACCGCGTCATGACC 327
QY 121 ATTGCCCTCGCTACGACCACTACGCTGACCGCTGGGACACCTCGGTGGTGGTGGACATC 180
Db 328 ATTGCCCTCGCTACGACCACTACGCTGACCGCTGGGACACCTCGGTGGTGGACATC 387
QY 181 ATGGGCGAGAAGTTCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATT 240
Db 388 ATGGGCGAGAAGTTCGCTACTTCCACTCCATCAAGAAGGGCGTGCACCGCGTGTGGATT 447
QY 241 GACCAACCGCTGTTCTCTGGCCAAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGGCCC 299
Db 448 GACCAACCGCTGTTCTCTGGCCAAAGGTCTGGGCAAGACCGGCTCCAAAGCTGTACGGCCC 506
```

Search completed: June 3, 2003, 15:27:21  
Job time : 1915.46 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 09:49:44 ; Search time 46.9794 Seconds  
(without alignments)  
12748.991 Million cell updates/sec

Title: US-09-980-771A-4

Perfect score: 1953

Sequence: 1 gcgcggacatcgtgatgt.....ccgccaccgcgaagtcgcgc 1953

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2.6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/prodata/1/ina/PCrus\_COMB.seq.\*  
6: /cgn2.6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475.8	24.4	2267	4	US-08-679-645-25
2	447.4	22.9	2542	3	US-08-941-445A-6
3	276.2	14.1	2067	4	US-09-388-743-21
4	266.6	13.7	2274	4	US-09-388-743-17
5	259.6	13.3	2176	4	US-09-388-743-13
6	255.4	13.1	2202	4	US-09-388-743-1
7	218.2	11.2	2825	4	US-09-196-390-5
8	216.8	11.1	2097	3	US-08-941-445A-10
9	210.8	10.8	2380	1	US-08-572-951-3
10	207	10.6	2248	4	US-09-345-214-20
11	206.6	10.6	1798	4	US-09-345-214-16
12	206.6	10.6	2019	4	US-09-345-214-15
13	161	8.2	2007	3	US-08-941-445A-8
14	161	8.2	2085	1	US-08-572-951-2
15	118.2	6.1	2239	4	US-09-196-390-1
16	109.6	5.6	2348	4	US-09-388-743-5
17	108	5.5	1758	3	US-08-836-567-3
18	106.8	5.5	2277	1	US-08-676-967-5
19	106.8	5.5	2277	1	US-08-676-974-5
20	106.8	5.5	2277	2	US-09-098-487-5
21	106.8	5.5	2418	4	US-09-388-743-25
22	106.6	5.5	2360	3	US-08-836-567-9
23	102.6	5.3	1528	4	US-09-345-214-6
24	102.6	5.3	1620	3	US-08-941-445A-20
25	102.6	5.3	1752	3	US-08-941-445A-12
26	102.6	5.3	2008	4	US-09-345-214-12
27	102.6	5.3	2491	4	US-09-345-214-5

28	102	5.2	2712	3	US-09-025-691-4	Sequence 4, Appli
29	101	5.2	2383	4	US-09-192-909-1	Sequence 1, Appli
c	99	5.1	4403765	4	US-09-103-840A-2	Sequence 2, Appli
31	97.8	5.0	2990	1	US-08-572-951-1	Sequence 1, Appli
32	97.4	5.0	4800	3	US-08-941-445A-4	Sequence 4, Appli
33	91.2	4.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
34	89.2	4.6	2793	3	US-08-836-567-7	Sequence 7, Appli
35	88.4	4.5	6027	2	US-08-968-542C-1	Sequence 1, Appli
36	86.8	4.4	2809	3	US-05-000-016-1	Sequence 1, Appli
37	86.8	4.4	2809	4	US-03-514-340-1	Sequence 1, Appli
38	86.4	4.4	44377	2	US-08-804-227C-7	Sequence 1, Appli
39	86.4	4.4	44377	2	US-08-804-198-1	Sequence 7, Appli
c	84.6	4.3	985	4	US-03-056-556-182	Sequence 1, Appli
c	84.6	4.3	985	4	US-09-072-596-177	Sequence 182, App
42	84.6	4.3	1620	2	US-08-461-775-10	Sequence 177, App
43	84.6	4.3	1620	3	US-09-031-606-10	Sequence 10, Appl
44	84.4	4.3	1926	3	US-08-836-567-5	Sequence 5, Appli
45	82.6	4.2	1464	1	US-07-735-065-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-08-679-645-25  
; Sequence 25, Application US/08679645  
; Patent No. 6350934  
; GENERAL INFORMATION:  
; APPLICANT: Zwick, Michael G.  
; APPLICANT: Edington, Brent E.  
; APPLICANT: McSwiggen, James A.  
; APPLICANT: Merlo, Patricia Ann Owens  
; APPLICANT: Guo, Lining  
; APPLICANT: Skokut, Thomas A.  
; APPLICANT: Young, Scott A.  
; APPLICANT: Folkerts, Otto  
; APPLICANT: Merlo, Donald J.  
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR  
; TITLE OF INVENTION: MODULATION OF GENE EXPRESSION  
; TITLE OF INVENTION: IN PLANTS  
; NUMBER OF SEQUENCES: 1263  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/679,645  
; FILING DATE: July 12, 1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/001,135  
; FILING DATE: July 13, 1995  
; APPLICATION NUMBER: 08/300,726  
; FILING DATE: September 2, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 219/247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELETYPE: 67-3510  
; INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 2267 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-679-645-25

Query Match 24.4%; Score 475.8; DB 4; Length 2267;  
Best Local Similarity 59.9%; Pred. No. 2.3e-74;  
Matches 979; Conservative 0; Mismatches 557; Indels 99; Gaps 7;  
QY 5 TGGACATCGTGTGCTGTGCTGAGTCCGCCCTTTGGTCCAAGACGGCGGCTGGGCG 64  
Db 418 TGAACGTCGCTTCGTGCGCGCGAGATGGCGCGTGGAGCAAGACCGCGGCTCGGCG 477  
QY 65 ATGTGACTGGTGGCTGCTATGTAGCTGGTCAAGCGCGGCCACCGCTCATCAACATTG 124  
Db 478 ACCTCCTCGCGCGCTTCCCGCGGCGCATGGCGCGAATGGCGACCGCTGCTGCTGT 537  
QY 125 CCGCTCGCTACGACCACTAGCTGACGCTGGACACCTCGGTGGTGGGACAT----- 179  
Db 538 CTCGCCGCTACGACCACTACAGGACGCGCTGGGACACCGCTGCTGCGGATCAAGA 597  
QY 180 -----CATGGCGAGAGGTCCGCTACTTCCATCCATCAAGAGGGGGTCCACGGCG 232  
Db 598 TGGGAGACAGGTACGAGACGGTCAAGTCTTCCACTGCTACAAAGCGCGGAGTGGACCGG 657  
QY 233 TGTGGATTCACACCCCTGCTTCTGCGCAAGTCTGGGCAAGACCGGCTCCCAAGCTGT 292  
Db 658 TGTGCTGTACCAACCCACTGCTTCTGGAGAGGTTTGGGAAAGACCGGAGGAGAATCT 717  
QY 293 ACGGCGCGCGCTCGGCGGTGACTACTGGACAACCAACAGCGCTTGGCCCTGTTCGCA 352  
Db 718 ACGGCGCTGACGCTGGAACGACTACAGGACCAACAGCTGCGGTTTCAAGCTGCTATGC 777  
QY 353 AGCGCGCTATGAGTGCOCGCGTGTGCTCCCTTGGG----- 390  
Db 778 AGGACGACTTGAAGTCCAAAGATCCTGAGCCTCAACAAACCCCATATCTTCCGGAC 837  
QY 391 --CCGCGGAGGACTGGTCTTCTGCGCAACGACTGGCACTCGCGCTGGTGGCGCTCC 448  
Db 838 CATACGGGAGGACTGCTGCTGCTGCAACGACTGGCACTCGCGCTTCTGCTGCT 897  
QY 449 TGTGAAGGACGATACCAAGCCCAAGGCGCAGTTTCAACCAAGGCCAAGTCGGTGTGCTA 508  
Db 898 ACCTCAAGAGCAACTACCACTCCACGGCATCTACAGGGACGCAAGACCGCTTTCGCA 957  
QY 509 TCACACATCGCTTCCAGGCGCGATGTGGAGGAGCTTTCAGGACACGAGCTGC 568  
Db 958 TCACACATCTCTTACGAGGCGGCTTCCGCTTCTCCGACTACCGGAGCTGAACCTCC 1017  
QY 569 CCGAGCGCGCTTTGACAAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGG 628  
Db 1018 CGGAGATTCAGTCTGCTTCTGATTTTCACTGACGCTACGAGAAG----- 1064  
QY 629 CCACCCCATGAGGAGGACGAGAACGCCCGCTGACGGGAAGACCTACAAGAGATCA 688  
Db 1065 -----CCGCTGGAAG-----GCCGGAAGATCA 1086  
QY 689 ACTGGCTGAAGGTTGGGATATTCGCGCGGACAGCTGGTGTGCTGCGCCCAACTACG 748  
Db 1087 ACTGGATGAAGGCGGATCTTGAAGCGGATGACATGAGGAGTGAACCCCAAGACGACA 1146  
QY 749 CGACCGAGATGCTTCCGCGGATCGCGCGGTGTGGAGCTGGACACCTCATCGCGCCA 808  
Db 1147 CCGAGAGCTCATCTCCGCGATCGCGAGGGCTCGGAGCTCGACAACATCATGCGCTCA 1206  
QY 809 AGGCATTTAGGGGATTTGAAGCGGATGACATGAGGAGTGAACCCCAAGACCGACA 868  
Db 1207 CCGCATCACCGGATCGTCAACGGCATGGACGTCAGGAGTGGGACCCCAAGAGGACA 1266  
QY 869 AGTTCCTGTGCGCTTACCAACAGAACAGCTTACGCGGCAAGCGCCCGCCCAAGG 928

Db 1267 AGTACATCCGCGTGAAGTAGGAGTGTGACGCGCGCTGGAGGCCAAGGCGCTCAACAAG 1326  
QY 929 AGGCCCTCAGAGCGGAGCTGGGCTGCTGTGTGACCCCAACCGCCCGCTTTCGCGCTCA 988  
Db 1327 AGCGCTCAGAGCGGAGGTCGGGCTCCCGGTGACCGGAACATCCCGCTGCTGGCGTTCA 1386  
QY 989 TCGGCGCGCTGGAGAGCAGAGGGTGTGGACATCATCTCTGGCGCGCGCTTCCCAAGATCC 1048  
Db 1387 TCGGCAAGGTTGAAGAGCAGAGGACCGGACGCTCATGGCGCGCGCATCCCGAGCTCA 1446  
QY 1049 TGGCCACCCCAAG---GTGCAGATCGCCATCTCTGGTACCGCAAGCGCGCTACGAGA 1105  
Db 1447 TGGAGATGGTGGAGAGCTGCAGATCTTCTGTGGGACGGGCAAGAAAGTTTCAGC 1506  
QY 1106 AGCTGTGAACGCGCATCGGACCAAGTACAAAGGCGCGCGCAAGGGCGTGGTCAAGTTCT 1165  
Db 1507 GCATGCTCATGAGCGCGGAGGAGAGTTCCAGGCAAGTGTGCGCGCTGTTCAAGTTCA 1566  
QY 1166 CGCGCGCGCTGGGCGACATGCTCACCGCGCGCGCGGACCTTCATGCTGCTGCTCGCGCT 1225  
Db 1567 ACGCGCGCTGGGCGCACCATCATGCGCGCGCGCGGACGCTGCTGCGCGTCAACAGCGCT 1626  
QY 1226 TCGAGCGCTCGCGCTGATCCAGCTGACGCGCATGTACTACGCTACGCTGCGCGCTGCTAG 1285  
Db 1627 TCGAGCGCTCGCGCTCATCCAGCTGAGGGATGCTATAGGAAACGCGCTTGGCGCTGCG 1686  
QY 1286 CTTCCACCGCGCGCTGCTGCTGACACCGCTCAAGGAGGCGCTCACCGGCTTCCACATGCGG 1345  
Db 1687 CGTCCACCGCTGGACTGCTGACACCATCATCGAAGGCAAGACGCGGTTCCACATGGGCG 1746  
QY 1346 CCGTGA-----ACCCGCAAGCTGACGAGCTGACGCGGAGCGGCTTGGCGCGCACCG 1399  
Db 1747 GCCTACGCTGCACTGCAAGCTGTGGAGCGCGGCGGACGCTCAAGAGTGGCGCACCT 1806  
QY 1400 TGGCGCGCTCAGGAGGCTGTTTGGCGCGCGCGCTTACCCCGAGATGTTGGCGCAACTGCA 1459  
Db 1807 TGCAGCGCGCTCAAGTGTGCGCACGCGCGGCTGACGAGGAGATGTTGAGGAACTGCA 1866  
QY 1460 TCAGCCAGGAGCTGCTGCTGTCGAAGCGCGCGCGCGCGGAGGAGGCTGCTGGAGGAGG 1519  
Db 1867 TGAATCCAGGATCTCTCTGGAAGGCGCTTCCAAAGACTGGGAGAGCTGCTGCTGCGCC 1926  
QY 1520 TGTGTAGCGCAAGGCGCGCTGCGCCACCGCAAGAGGAGGAGATCAAGGTGCGCGCTTG 1579  
Db 1927 TCGGGGTGCGC---GGCGCGGACGAGGCTGGAAGCGGAGAGATCGCGCGCTCGCCA 1983  
QY 1580 CCGAGAAGATCCCGC 1594  
Db 1984 AGGAGAAGTGGCGC 1998

RESULT 2

US-08-941-445A-6

Sequence 6, Application US/08941445A

Patent No. 6107060

GENERAL INFORMATION:

APPLICANT: Keeling, Peter

APPLICANT: Guan, Hanping

TITLE OF INVENTION: Starch Encapsulation

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.

STREET: 5370 Manhattan Circle

CITY: Boulder

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:



```

: APPLICATION NUMBER: US/08/941,445A
: FILING DATE: 30-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,855
: FILING DATE: 30-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 89-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2542 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 453..2282
: US-08-941-445A-6

```

[illegible]

Db	1300	CCGAGAGGTTCAAGTTCATCCTTCGATTCATTCACGCGGATGAC-----	1343
QY	629	CCACC CCCATGGAGGAGGACGAGAAGCCCGCTGACGGGAAAGACCTACAAGAAGATCA	688
Db	1344	-----ACGCGGTGGAGGGCAGGAAGATCA	1368
QY	689	ACTGGCTGAAGGTGGCAATTATCCGCCGCCACAAAGCTGTGACTGTCTGCCCAACTACG	748
Db	1369	ACTGATGAAGGCCGGAATCTCTGGAAGCCACAGGGTGCTCACCGTGAGCCGCTACTACG	1428
QY	749	CGACCAGAGTCGCTCCGCGATCCGCCGGGTGTGAGCTGGACACGCTCATCCGCGCCA	808
Db	1429	CCGAGGAGCTCATCTCCGSCATCCCAAGGGATCGAGCTCGACACATCATCGCGGTCA	1488
QY	809	AGGGCAATTGAGGCAATTGTAAGCGCATGGACATTTAGGAGTGGAAACCCCAAGACCGGACA	868
Db	1489	CCGGCATCACCGGCATCGCTCAACGGCATGGACCTCAGCGAGTGGATCTCTAGCAAGGACA	1548
QY	869	AGTTCCTGCTCTCGCCCTACGACCAACAGCGTCTACGCCGCGAAGGCCCGCCCAAGG	928
Db	1549	AGTACATCACCGCAAGTACGACGCAACCGCAATCGAGGCGAAGGGCGCTGAACAAGG	1608
QY	929	AGGCCCTGCAAGCCGAGCTGGGCCCTGCCTCTGGACCCCAACGCCGCCCTGTTCGCGCTTCA	988
Db	1609	AGGCGTTGCAAGCGGAGGCGGTCTTCGCGTGCACAGGAAATCCCACTGATCGGTTCA	1668
QY	989	TCGGCCGCTTGAGGAGCAGAGGGTGTGGACATCATCTCTGCCGCCCTGCCCAAGATCC	1048
Db	1669	TCGGCAGGCTGGAGGAACAGAGGGCCCTGAGCTCATGCCGCCGCATCCCGAGACTCA	1728
QY	1049	TGGCCACCCCAAGGTGCAGATCGCCATCTGGGTACCGGCAAGGCCGCTACGAGAAAGC	1108
Db	1729	TG----CAGSAGGACGCTCCAGATCGTTCTTCTGGTACTGGAAGAGAAGATTCGAGAAGC	1785
QY	1109	TGTTGAACGCCATCGGCACCAAGTACAGGGCCGCCCAAGGGCGTGGTCAAGTCTCGG	1168
Db	1786	TGCTCAAGAGCATGAGGAGAAGTATCCGGGAAGGTGAGGCGGTGGTGAATTTCAAG	1845
QY	1169	CGCCCTTGCGGCACATGCTCACCGCGCGCGACATTCATGTTGGTGCCTCGCGGTTTG	1228
Db	1846	CGCCGCTTGCTATCATCTATGCGCGGAGCCGACGTGCTGCGGCTCCCAAGCGGTTTG	1905
QY	1229	AGCCCTCGGCCCTGATTCAGCTGTCACGCGCATCACTACGCTACCGGTGCCGTGGTAGCCT	1288
Db	1906	AGCCCTGTGGACTCATCAGCTGCAAGGGGATGAGATACGGAACGCCCTGTGCTGGCGCT	1965
QY	1289	CCACGGCGGCTGTGTCACACCGTCAAGGAGGCGTCAACGGCTTCACATGGGGGCC	1348
Db	1966	CCACGGTGGGCTGTGGACACGGTCTCAGAGGCAAGACTGTTTCCACATGGGCGGCTC	2025
QY	1349	TGAACCCCGAC-----AAGCTGGACAGGCTGAGCGCGACGCCCTGGCCGACCGCTGC	1402
Db	2026	TCAGCGTCCGACTGCAAGTGTGTGGAGCCAAAGCGCATGAAGAAGGTGGCGGCCACCTGA	2085
QY	1403	GCCGTGCCAGCGAGGTGTTGCGGGCGGCCGTACCCCGAGATGTGGCGCAACTGCATCA	1462
Db	2086	AGCGGCCATCAAGTCTGTGCGGACCGCGGGGTACGAGGAGATGTTGAGAACTGCATGA	2145
QY	1463	GCCAGGACCTCTCTGTGTCGAAGCCGCCCAAGTGGAGGGCCTGCTGGAGGAGGTGG	1522
Db	2146	ACCAGGACCTCTCTTGGAAGGGCCCTGGAGAACTGGGAGAAATGTGCTCTGGGCCCTGG	2205
QY	1523	TGTACGCAAGGGGCGGTGGCCACCG	1549
Db	2206	GCCTGCGCGGACGCGCGCGGGGATCG	2232

RESULT 3  
US-09-388-743-21  
; Sequence 21, Application US/09388743  
; Patent No. 6423886  
; GENERAL INFORMATION:

```

; APPLICANT: Singletary, George
; APPLICANT: Zhou, Ian
; TITLE OF INVENTION: No. 6423886a1 Starch Synthase Polynucleotides and Their
; TITLE OF INVENTION: Use in the Production of New Starches
; FILE REFERENCE: 1144
; CURRENT APPLICATION NUMBER: US/09/388,743
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Tulipa fosteriana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58)...(1857)
; US-09-388-743-21

```

Query Match	14.1%	Score 276.2	DB 4	Length 2067
Best Local Similarity	51.9%	Pred. No. 7.5e-40		
Matches 820	Conservative 0	Mismatches 663	Indels 96	Gaps 5
QY 1	CGCTGGACATCGTATGTTGCTGCTGAGTGCGCCCTTGCTCAACAGCGGGCGCCTG	60		
DB 271	GGGATGAACCTGGTGTTCGTGGGACCGAGAGGTCGTCAGCAAGACCGGGGGTC	330		
QY 61	GGCGATGTACTGTGGCTCCCTATTGAGCTGTCAAGCGCGGCCACCGGTCTAAGCC	120		
DB 331	GGGGATGTCTAGGAGGGTTACCGCCGCCCTGGCGGAGAGGGCATCGGGTCAIGTT	390		
QY 121	ATTGCCCTCTCTAGGACCACTAGCTGACGCTGGGACACCTCGTGGTGTGACATC	180		
DB 391	GTCACTCCGCGGTACGATACAGGATGATGGGACAAACACTGTGCTTGAGATC	450		
QY 181	ATGGSC-----GAGAAGTCCGCTACTTCCACTCCATCAAGAAAGGCGCTGAC	228		
DB 451	AAAGTCGGGGATAAAATGAGACGGTTCGTTCTCCACCTCCATAGAGGGGGTGGAT	510		
QY 229	CGCGTGTGGATTGACACCCCTGGTTCCTGGCCAAAGTCTGGGGCAAGACCGCTCCAAG	288		
DB 511	AGSGTGTATTGATCACCTTGGTTCCTGAGAAAGTTGGGGGAAACCGGTGGGAA	570		
QY 289	CTGTACGCCCCCGCTCCGGGCTCACTACTGGCAACCCCAAGCGCTTCGCCCTGTTC	348		
DB 571	TTGTATGTCTCTTTACTTGGAACTGATTATGATGATACCACTAGGGTTCAGTCTTCTG	630		
QY 349	TGCAAGCGCGCTATTGAGGCTGCCCGCTGTGCCCTTCGGCCCC-----	393		
DB 631	TGTCAGGCTGCTTGGAGGCTCCAAGAGTTCATAATCAACAACAGTAAATATTTCT	690		
QY 394	-----GGCGAGGATGCTCTTCTGGGGCCACGACTGGCACTCCGCCCTGGTCCCC	444		
DB 691	GSACCATATGTTGAAGATGCTGTTTTATTGGGAACGATGGCACTGGACTCTTCCA	750		
QY 445	GTCTGCTGAAGACGAGTACCAGCCCAAGGGCCAGTTCACCAAGGCCAAGTCGGTGTG	504		
DB 751	TGCTACTGAAGAGTGTATAAATCAAGAGGATTATGAGAGTGCCTAAGGTTCCTTT	810		
QY 505	GCTATCCACACATCGCTTCCAGGGCCGCGCATGTGGGAGGAGGCTTTCAGGACACGAA	564		
DB 811	TGCATTCAATATGGCATACCAAGGCAGATTCGCTTCCTGATTTCTCGCTTCTCAAC	870		
QY 565	CTGCCCCAGCGCCCTTTTCAAGAGTGGCTTCTCGGACGGCTATGCCAAGGTTTACACT	624		
DB 871	CTTCCAGACACATTTAAATTCCTGTTTGAATTTCTTCGATGATATACAAACCT-----	924		
QY 625	GAGGCCACCCCATGGAGGAGGACGAGAGACGCCCGCTGACGGGAAAGACCTTCAAGA	684		
DB 925	-----CTGAAGGTAGAGAA 939			
QY 685	ATCAACTGCTGAAGGTTGGCATTTATCGCGCGCACAAAGCTGGTGACTGTGTCGCCAAC	744		
DB 940	ATAAATTGATGAAGCTGGAATATTGGAAGCGGACACTGTGTACTGTGAGCCCGTAT	999		

; LENGTH: 2274  
; TYPE: DNA  
; ORGANISM: Typha latifolia  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (115)...(1956)  
US-09-388-743-17

Query Match 13.7%; Score 266.6; DB 4; Length 2274;  
Best Local Similarity 52.9%; Pred. No. 3.5e-38;  
Matches 828; Conservative 0; Mismatches 639; Indels 98; Gaps 8;

QY 1 GCGTGGACATCGTATGGTTGCTGCTGAGTGGCCCTTGGTCCAGACGGGGCGCTG 60  
Db 373 GGGATGAACCTAGTCTTTGGAGCTGAGATGGCTCCATGGACGACACCTTGGGCCCT 432  
QY 61 GGCATGTGACTGTGGCTGCTATTTAGCTGGTCAAGCGCGGCCACCGGCTCATGCC 120  
Db 433 GGTGATGTTCTTGGAGACTCCACCGGCATTTGGCGCAATGGACATCGAGTTATGGTT 492  
QY 121 ATTGCCCTCGCTACGACCATGCTGACGCTTGGGACACTCGGTGGTGGTGGTGGTCAAT- 179  
Db 493 ATACGCCACGTTATGATCAATATGATGATGCTTGGGATACAGATGCTCTTTGAGTTG 552  
QY 180 -----CATGGCGGAGAGTCCGCTACTTCCACTCCATCAAGAAGGGCGTGAC 228  
Db 553 AAAGTTGGGATAGTGTGAACCGTGGCTCTTCTTCACTGCTATTAAGAGGAGTTGAT 612  
QY 229 CGGTGTGATGACACCCCTGTTCTTGGCCAAAGTCTGGGCAAGACCGGCTCCAAAG 288  
Db 613 CGAGTTTGTGATCAACCTATGTTCTTGGGAAGTCTGGGGAACACTGGTGGGAAG 672  
QY 289 CTGTACGGCCCGCTCGCGGCTGACTACTGTGACACCAACACAGCGCTTGGCCCTGTT 348  
Db 673 ATTATGGTCTTACACTGGAACAGACTATGAGCAATACAGTACGCTTACAGCTTTCTA 732  
QY 349 TGAAGCGCTATTTAGGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 393  
Db 733 TGGCAGGACGACTTGAAGCTCTAGAAATCTTAATCTCAACACAGTATCTTCTCT 792  
QY 394 -----GGCGAGGACTCGCTTCTGCTGGCCAAAGCTGGCACTCGGCCCTGGTCC 444  
Db 793 GGTCTTAAAGGAGATGTTATCTTCAATTTGCAATGATTGGCACACTTCTTCTGCCA 852  
QY 445 GTCTGTCTGAAGGACGAGTACCAAGCCCAAGGCGGAGTTCACCAAGCCAGTCTGCTG 504  
Db 853 TGCTACTTAAAGAGATGTATACATCCCGCTGGCATTTACAAAGACCGCAAGTTGCTTC 912  
QY 505 GCTATCCACAACTGCTTCCAGGCGCGCATGTGGAGGAGGCTTTCAAGGACACGAAG 564  
Db 913 TGCATTCACAATATATACACCAAGTTCGATTTCTCCCTCAGACTTCGAATTTCTCAAT 972  
QY 565 CTGCCCGAGCGCCCTTTGACAGCTGGCTTCTCGACGGCTATGCCAAGGTTTACACT 624  
Db 973 CTTCGGAG-----AATTCAAATCT 993  
QY 625 GAGGCGACCCCATGAGGAGGACGAGAAAGCCCGCTGACGGGAAAGACCTACAAGAG 684  
Db 994 TCTTTTCAGTTTCATGATGGTACACAGCC-----TGTGAAGGAATGAAG 1041  
QY 685 ATCACTGCTCAAGGTTGGCATTTATCGCGCGCAGCAAGCTGTGTGACTGTGCGCCCAAC 744  
Db 1042 ATAAATTTGATGAAGCAGGAATCTTAGAATCAGACAGGTTTTCAGTGAGTCCATAT 1101  
QY 745 TAGCGACCGAGATCGCTGCCGATCGCGCGGCTGTGGTGGTGGTGGTGGTGGTGGTGG 804  
Db 1102 TAGCACAAGAGCTCTTTAGGAGAAGAAAGGGAGTCTGAGTGGACACATTTTGGCT 1161  
QY 805 GCGAAGGCGATTGAGGCGATTGTGAACGCGCATGGACATTTGAGAGTGAACCCAGAC 864  
Db 1162 GTGACCAAGTATCAGGAATTTGATGGAATGGATGTTAATGAGTGGAAATCCATTAACA 1221  
QY 865 GACAAGTTCTCTGCTGGCGCTTACGACCAAGACAGCTCTACCGGCGCAAGCGCGCGC 924

Db 1222 GACAAATATATTTCTGTTAATATGATGCAAAACCTGTATGTAAGCAAGGCTCTTAAC 1281  
QY 925 AAGGAGGCGCTCGAGGCGG-AGCTGGCGCTGCTGGGACCCACCGCCCTCTGTCCG 983  
Db 1282 AAGGAGCATTCGAAGCTGAAAGTTGGCTTGGCTGTAGACAGGACATCCCTGTATTTG 1341  
QY 984 CTTTCATCGCCCGCTTGGAGGAGCAGAAAGGTTGTGACATCATCTTGGCGCGCTGCCAA 1043  
Db 1342 ATTTCAT-GGAAGCTAGAGGACGACAAAGATCATATTTAGCAGCATTCATTCAG 1400  
QY 1044 GATCTGCGCCCGCCCAAGGTGACATCCCATCTGGGTACCGCAAGCGCGCTTACGA 1103  
Db 1401 GATCATGG---ATGAGAATTTTCACTAATTTCTTGGAACTGCAAGAAGAAATGA 1457  
QY 1104 GAAGTGTGTGAACGCTCGGCAACCAAGTACAAAGGCGCGCCCAAGGGGTGGTCAAGTT 1163  
Db 1458 GAATCAGTTGAGAGTATGGAGGAATGTTCGCCGACAAAGTGTGGGCGAGTTATGAAGTT 1517  
QY 1164 CTCGCGCGCTTGGGCGACATGCTCACCGCGCGCGCGCTTCACTGCTGGTGCCTCGG 1223  
Db 1518 CAATGCTCCCTTAGCTCACCAGATGACGCGGGAGCTGATATAATTTGATTCATTC 1577  
QY 1224 CTTGAGCGCTTGGCGCTGATCCAGCTGCACGCCATGACGTACCGTACCGCTGGT 1283  
Db 1578 ATTGCAACCATCGCGCTTATCCAGTTGCAAGGCTCAATATGAAAGCGCTTCTGCGTG 1637  
QY 1284 AGCTCCACCGCGCGCTGTCGACCGCTCAAGGAGGCGCTCACCAGCTTCCATGCG 1343  
Db 1638 TTCTCCACTGCTGGCTTGGGACACGCTGAAAGAGGCAAACTGGATTCATATGG 1697  
QY 1344 GCGCCTGA-----ACCCGACAAAGCTGGACGAGCTGACCGCGCTGCGCGCGCAC 1397  
Db 1698 ACCTTTCAGTCTGAATCGGAAGTGGTCTTCTTCTGATGAAGAAAGTTGTCACAA 1757  
QY 1398 CGTGGCGCTGCGCAGCAGGTTGTTGGCGCGCGCTTACCCCGAGATGTTGGCGCACTG 1457  
Db 1758 TGTGAAAGCGGCACTTAAGGTCATCGGCACACCTTGCCTTTGAGGATATGATCAAGAATTG 1817  
QY 1458 CATCAGCAGGACCTCTCTGTCCTCAAGCGCGCGCGCTGACCGCGCTGCGCGCGCAC 1517  
Db 1818 CATGGCACAGACCTCTCATGGAAGGCGCTGCAAAAGACTGGGAGCAAGTTCTACTGAA 1877  
QY 1518 GGTGG 1522  
Db 1878 CTTGG 1882

## RESULT 5

US-09-388-743-13  
; Sequence 13, Application US/09388743  
; Patent No. 6423886  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their  
; FILE REFERENCE: 1144  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 2176  
; TYPE: DNA  
; ORGANISM: Canna edulis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (81)...(1928)  
US-09-388-743-13

Query Match 13.3%; Score 259.6; DB 4; Length 2176;  
Best Local Similarity 51.6%; Pred. No. 5.6e-37;

Matches	795;	Conservative	0;	Mismatches	649;	Indels	96;	Gaps	57;
5	QY	TGGACATCGT	GATGGTTGCTCTGAGTGC	CGCCCTTGGTCC	CAAGACGGCGCCCTGGCG	64			
349	Ddb	TGAACTTGGT	TTGGTTGTGAGGTAGCTCC	ATGGAGCAAACTGGGGCC	TTGGCG	408			
65	QY	ATGTGACTGGT	GGCTGCCTATTGAGCTGT	CAAGCGCGCCACCCGCTCATG	ACCATTG	124			
409	Ddb	ATGTTCTTTAG	AGGATGTCACCTGCTATGGCTGCA	ATTGGGCACAGGGTCATG	ACCGTGG	468			
125	QY	CCCTCGCTAC	GCACAGTACGCTGACCGCTGG	GCACACCTCGTGGTCTGG	GCACAT	179			
469	Ddb	TCCACGATATG	ACCAATATAAGATATCTGG	GTACAAAGTGTCC	CAGTTGAGTTAAAG	528			
180	QY	-----CAT	GGGCGAGAAGTCCGCTACTT	CCATCCATCA	AGAGGGGGTGCACCGC	232			
529	Ddb	TTGGGGAT	TAAGATTGAAACTGTCGCTCTTCC	ACTGTCTACAAAGGGGAGTTG	ATCGGG	588			
233	QY	TCGTGATTG	ACACCCCTGTTCTGCGCAAGGTCT	GGGCAAGACCGGCTCCA	AGCTGT	292			
589	Ddb	TTTTTGTG	GATCACCCCTATGTTTCTC	AGRAGGTTTGGGGAA	ACAAGGAGAAATTAT	648			
293	QY	AGGCCGCCCT	TCGGGCGCTGACTACTCTG	GCACCAACCAAGCGGTT	CGCCCTGTTCTGCA	352			
649	Ddb	ATGTCCTCTT	TACAGGAACAGATTATG	CAGACAATCAACTAAG	ATTTCAGCTTTTGTC	708			
353	QY	AGGCCGCTAT	TGAGCTGCCCGCTGCTCCCTT	CGCCCGCGGAGGATGC	-----	405			
709	Ddb	TGGAGCTCTG	GAAGCTCCAGACTTCAAGACTT	CAATCAACACAGCAAT	ACTATCTTGGAC	768			
406	QY	-----GT	CTTCGTGGCCAAAGACTGG	CACTCGCCCTGGTGC	CGCGTCC	448			
769	Ddb	CATATGG	AGATGATGTTGTGTTTATTG	CAACAGATTGSCATCT	CTCTACTGCGCTGCT	828			
449	QY	TCGTAAG	ACGAGTACACGCCCAAGGCC	AGTTTCCACAGGCCAAGT	CTGCTGCTA	508			
829	Ddb	ACTTTGAAA	ACTATGTACCAATCACTAGTATTT	ACATGAATGCTAAGGTTG	CAATTTGCA	888			
509	QY	TCACAACA	TATGCCCTTCAGGGCCG	CACTGTGGGAGGAGCTT	TCAAGACACGAAGTGC	568			
889	Ddb	TTTCATAT	ATTGCTTACCAGGGCCG	GATTTCGCTTTCCG	ACTTTTGAATCTCTAATCTCC	948			
569	QY	CCAGCGCG	CGCTTTGACAAAGCTGG	CGCTTCTCGAGCGCTAT	CGCAAGTTTACACTAGG	628			
949	Ddb	CCAATAA	ATTTAAATCTTCAITTTG	ATTTCAITGGATGGAATAT	GACAAACCT-----	998			
629	QY	CCACCCCAT	GAGGAGGACGAGAGCCCG	CTGACGGGAAAGACCTT	ACAAGAAGATCA	688			
999	Ddb	-----GT	AAAGGAAGAAATAA	-----	-----	1017			
689	QY	ACTGGCTG	AAGGTTGGCATTTATCG	CGCCGACAGCTGGT	GACTGTGTGCGCCCACTACG	748			
1018	Ddb	ATTGGAT	GAAGCTGGAATAATAGA	ATGTGATGTTGCTT	GACCGTGAGCCCATATTATG	1077			
749	QY	CGACCGAG	ATCGCTGCCGATGCGCG	GGGTGTGGACTGGAC	ACCGCTATCGCGCCA	808			
1078	Ddb	CCAAGAG	CTGTCTCAGGGGTAG	AGAAGGGTTGAGTTGGG	CAATCTCTGCGCATGA	1137			
809	QY	AGGGCAT	TGAGGCAATGTGA	CGGCATGGCAATTTGAG	GAGTGGAAACCCCAAGACCGACA	868			
1138	Ddb	AAACCAT	CTCTGAAATAGTAAAT	TGGATGGACACACCG	AGTGGAAATCCATTAA	1197			
869	QY	AGTTCTGT	CTGGCCCTACGACAC	AGACAGCGTCTAC	CGCGGCAAGGCCCGCGCAAGG	928			
1198	Ddb	AATATAT	TTCTACAAACTACG	ATGAACTGTATTTGG	ATGCAAAACCTCTCTGTAAAG	1257			
929	QY	AGGCCCTG	CAGGCCGAGCTGG	CGCTTGTGGACCC	CCACCGCCCTGTTTGGCCTTCA	988			
1258	Ddb	AAGCTT	GTCAAGCTGAGTGTGG	CGCTGCCCTGTAA	CAAAAACAAGCTTGT	1317			
989	QY	TCGGCG	CGCTTGGAGGACGA	AGGGTGTGGACAT	CATCCTGGCGCCCTGCCAAGATCC	1048			
1318	Ddb	TTGGAAG	ACTTAGATGAGCA	AGAAAGGCTCAGACAT	TCTAGCTGCAGCAATTCGAGACATC	1377			

QY 1049 TGCCACCCCAAGCTGCAGATCGCCATCTCTGGGTACCGCAAGCGCCCTACGAGAAGC 1100  
 Db 1378 T----TTGTGAGAAATGTTCAAGTGATAGTACTTTGGCACTTGGCAAGAGAGTTGGAGATG 1434  
 QY 1109 TGGTGAACGCCATCGGCACCAAGTACAAGGGCCGCGCAAGGGCGTGGTCAAGTTCCTGG 1168  
 Db 1435 AACTTACATTACTTTGAGGAAATGTTTCCAGACAAATTCAGAGCACATCTCAAAATTCACG 1494  
 QY 1169 CGCCCTCGGGGCACATGCTCACCGCGCGCGGACTTCATGCTGTGCTCCCTCGCGCTTCG 1228  
 Db 1495 TTCTTTAGCTCATGCAATCATGGCAGGAGCTGATATCCTTGTATTCCAAAGCAGATCG 1554  
 QY 1229 AGCCCTCGGCCCTGATCCAGCTGCACGCCATGCATCTAGCTAGTACCGTGCCTGGTAGCGCT 1288  
 Db 1555 AACCTGTGCCCTCATTCAGCTTCAGGCCATCGGATATGGAATCTCCCTATGTGTAGCA 1614  
 QY 1289 CCACCGCGCGCTGTGCACACCGTCAAGGAGGGCGGTACCGCGCTTCCACATGGCGCC- 1347  
 Db 1615 CCATGTGTGACTTTGTGACACTGTCAAAAGAGGGCTTCACTGGCTTCCATATGGGCCCT 1674  
 QY 1348 ----CTGAACCCCGACAAGCTGGACGAGGCTGAC3CGACGCCCTGGCCGCCACCGTGC 1402  
 Db 1675 TCAGTGTGGAGTGTATGCCGTAGACAAGCTGATCTACAAAAGATTTCTGAAACACGA 1734  
 QY 1403 GCGGTGCCAGGAGGTGTTTGGCGGGCGCGGCTACCCGAGATGTGGCCAACTGCATCA 1462  
 Db 1735 AAAGGCCCTCAAAGTCTATGTAACACCTGCTTTTGGAGATGATCAAGAACTGCATGA 1794  
 QY 1463 GCAGGAGCTGTCTGTGTCACGCCCGCCCAAGAGTGGGA 1502  
 Db 1795 ACCAAGATCTCTATGAAAGGACCTGCAAGAGAGTGGGA 1834  
 RESULT 6  
 US-09-388-743-1  
 ; Sequence 1, Application US/09388743  
 ; Patent No. 6423886  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Singletary, George  
 ; APPLICANT: Zhou, Lan  
 ; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and  
 ; TITLE OF INVENTION: Use in the Production of New Starches  
 ; FILE REFERENCE: 1144  
 ; CURRENT APPLICATION NUMBER: US/09/388,743  
 ; CURRENT FILING DATE: 1999-09-02  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2202  
 ; TYPE: DNA  
 ; ORGANISM: Curcuma zedoaria  
 ; NAME:  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (130)...(1974)  
 US-09-388-743-1  
 Query Match 13.1%; Score 255.4; DB 4; Length 2202;  
 Best Local Similarity 51.8%; Pred.No.3e-36;  
 Matches 799; Conservative 0; Mismatches 646; Indels 96; Gaps  
 QY 5 TGGACATCGTGATGGTTGCTGCTGAGGTGCGCCCTTGGTCCAAAGCGGCGGCTGGGGG 64  
 Db 395 TGAACCTTGATCTTTTGTGCTGCTCAGGTGGCTCCGTGGAGTAAACTGGAGGCTTGGTG 454  
 QY 65 ATGTGACTTGGTGGCTGCTGCTATTGAGCTGGTTCAGCGCGGCCACCGCTCATGACCATTTG 124  
 Db 455 ATGTTCTTTGGAGGTTTGGCCACCGCCATTTGGCCGCGGCGGCAAGGACACAGGTCATGACTATAG 514  
 QY 125 CCCCPCGCTACGACACAGTACGCTGACGCTGGGACACCTCGGTGGTCTGGAGCATCATGG 184  
 Db 515 GACCGGACATGACCAATACAAAGATGATGGATACGGCTGCTTTGTGAGTTGAGGAG 574

Query Match	Score 255.4; DB 4; Length 2202;
Best Local Similarity 51.8%;	Pred. No. 3e-36;
Number of matches 645;	Indels 96;
Number of gaps 6	

QY	5	TGGACATCGTGATGGTTGCTGTGAGGTGCCCCCTTGGTCCAGACGGCGCCCTGGGG	64
Db	395	TGAACCTTGATCTTTGTTGCTGTGAGGTGGCTCCGTGGAGTAAACTTGAGGGCTTGGT	454
QY	65	ATGTGACTGTGGCGCTGCCATTAGCTGTCTAAGCGCGGCCACCGCGTCTATGACCATTTG	124
Db	455	ATGTTCTTGGAGGTTTGCCACCGGCCCATGCGGCAAGGGGATGATGACTATAG	514
QY	125	CCCTCGCTACGACCACTACGCTGACGCCCTGGGACACCTCGTGGTCTGTGGACATCATGG	184
Db	515	CACCGGCACATGACCCATATCAAGATGGATGGCATACGGCTGCTCTTTGTCAGCTTGAAG	574

QY	185	-----GCGAGAAGGTCGGCTACTTCCACTCCATCAAGAAGGCGTGCACCGG	233
Db	575	TTGTGTATAGAATTTAAACTGTTTCGCTTTTTCACCTCCACAAAAGGAGTTGATCGG	634
QY	233	TGTGGATTGACCAACCCCTGGTTCCTGGCCAAAGCTCTGGGGCAAGACCGGCTCCAAAGCTGT	292
Db	635	TGTTTGTGGATCACCCTCTCTTCTCTTGAGAAGGTTTGGGGAAGAACTGAGGAAAGATAT	694
QY	293	ACGGCCCCGCTCCGGCGCTGACTACCTGTGACAAACACAAAGCGCTTCGCCCTGTCTCGCA	352
Db	695	ATGGTCTGTGCACAAGAACTGATATGAAGACAACCAAGCTAAAGTTTCTGTCTCTGTGTT	754
QY	353	AGGCGCGTATTGAGGCTGCCCGGTGCTG-----CCCTTCGGCC	391
Db	755	TGGCAACTCTGGAACATCCAAGGTTCTGAAATCCCAACAATACAAATATCATTTCTGGAC	814
QY	392	CCGCGGAGGACTGGTCTTCGTGGCCAAACGACTTGGCACTTCGGCCCTGGTGGCCGCTCTGC	451
Db	815	CAAAAGTGGAAGATTATTCATTTCTTAACGATTGGCATACTGCTCTATTACCTTGCTATT	878
QY	452	TGAAGGAC---GAGTACCAAGCCCAAGGCGAGTTTCAACCAAGCCCAAGTCGGTCTGGCTA	508
Db	875	TAAAGACCATGTTATATCAAGCCCATGGAATATACAAAATGCTAAAGTTGTTTCTGCA	934
QY	509	TCCACAACATCGCCTTCCAGGGCCGCAATGTGGAGAGAGCTTTCAAGGACACGAAGCTGC	568
Db	935	TTCTAATAATTGGGTATCAAGGACGGTTTGCCCTTTGAAGATTTTTCGGGTCTCAATCTCC	994
QY	569	CCAGGCGCGCTTTTGACAAGCTTGGCCCTTCTCGSACGGCTATGCCAAGGTTTACATGAGG	628
Db	995	CTGNTACATTTCAAGTCTCTTTTGATTTTCATCGATGGCTATGCAAA-----	1040
QY	629	CCACCCCATCTGAGGAGGACGACGAAGCCCGCTGACGGGAAGACCTACAGAAGATCA	688
Db	1041	-----ACCAATAAAGAGAGAAATCA	1063
QY	689	ACTGGCTGAAGGTGGCATTTATCGCCGCCGACAAGCTGGTGACTGTGTGCCCAACTACG	748
Db	1064	ACTGGATTGAAGCGGGAATTTAGAAATCAGATCGTGCATTGCTGTGAGCCCACTATG	1123
QY	749	CGACCGAGATCGCTGCGGATGCGCGCGGTGTGGAGCTGGACACCTCATCCGCGCA	808
Db	1124	CCAGGAACTGCTCAGGAATCGATAAGGGCGTCGAGTTGGACAATATACTGGCGCTTGA	1183
QY	809	AGGGCATTTGAGGCAATTGTGAACGGCATGGCACTTGAGGAGTGGAAACCCCAAGACCGACA	868
Db	1184	AAACCATCTGTGCGCATCATAAATGGAAATGACACCAACGAGTGGAAATCCCTCACAGACA	1243
QY	869	AGTTCTGTCTGCGCCCTACGACCAAGACAGGCTCTACCGGGCAAGCGCGCCCAAGG	928
Db	1244	AATACATAACAGAAATTTACGACCAACCACTGTAAATGGAGCAAGGCCACTCAACAAGG	1303
QY	929	AGGCCCTGCAAGCGGAGCTGGCGCTCGCTGTGGACCCCAACCGCCCTGTTCGCGCTTCA	988
Db	1304	AAGCTTTGCAACTGAGTTGACTTGGCTGCGCTGCAACAGTAAATCCCTGTGTAAGCTTTCA	1363
QY	989	TCGGCCGCTTGGAGGACGAAGAGGTGTGGACATCATCTCGCGCGCCCTGCCCAAGATCC	1048
Db	1364	TTGGCAGACTAGAAGAACAAAAGGTTTCAGACATTTCTAGCTGAAGCAATTTCCAAGTTCT	1423
QY	1049	TGSCCAACCCCAAGTGTGAGATGCGCCATCTGGGTACCGGCAAGCGCGCCCTACGAGAAGC	1108
Db	1424	T---CGATCAGGATGTTCAAGTGCATAGTTCTCGGTACTTGGTAAAAAGAAAGTTAGAGCGCC	1480
QY	1109	TGCTGAACGCCATCGSCACCAAGTACAAGGCGCGCCCAAGGCGTGGTCAAGTTCTCGG	1168
Db	1481	AACTTGCATTGCTCGAGGACGAGTTTCCACAGACAAATTCAGAGCTCATATGAAGTTCAATA	1540
QY	1169	CGCCCTTGGCCACATGCTCAACCGCGCGCGGCTTCATGCTGGTGCCTTCGCGGCTTCG	1228
Db	1541	TTCCCTTTGGCTCATGGAATCATGGCGGGTGTGATACCTTGTATTCCAGTAGGTTTCG	1600
QY	1229	AGCCCTGCGGCGCTGATCCAGCTGCAAGCCCATGCACTACGGTACGGTACGGTGCCTGTAGCCT	1288

Db	1601	AACCATCGGTCTATTTCAGTCTCCAGGCATGAGATACGGAACCCCTTCCATGTGCACCA	1660
Qy	1289	CCACCGCGGCCCTGGTGCACACCGTCAAGGAGGGGCTCACCGGCTTCCACATGGCGCC-	1347
Db	1661	CGACTGTGTGGGCTCGTCGACACTGTCAAAAGAGGCATCACAGGTTTCCACATGGTCCCT	1720
Qy	1348	-----CTGAACCCCGCAACAGCTGGAGGGGCTGACGCGAGCCCTGSCCGCCACCCGTGC	1402
Db	1721	TCAGTGTGGAGTGCAGACATGCCGACGAGCGCAGCTCTAAAGATGTGGAACCATGA	1780
Qy	1403	GCGTGCACGAGGHTTTGGGGGGGCGCTACCCGAGATGGTGGGCCAACTGCATCA	1462
Db	1781	AGAGAGCCCTTATGGTTATGGAACGCTGTCTTCGAGGAGATGATACAGAAGTGCATGG	1840
Qy	1463	GCGAGCACTGTCTCGTCCAAAGCCCGCCCAAGTGGGAG	1503
Db	1841	CTCAAGATTCTCTGTAAGGGGCCAGCAAGGAATGGAG	1881

## RESULT 7

US-09-196-390-5  
: Sequence 5, Application US/09196390  
: Patent No. 6307125  
: GENERAL INFORMATION:  
: APPLICANT: Block, Martina  
: APPLICANT: Lorz, Horst  
: APPLICANT: Luticke, Stephanie  
: APPLICANT: Walter, Lennart  
: APPLICANT: Froberg, Claus  
: APPLICANT: Kossmann, Jens  
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
: TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
: TITLE OF INVENTION: SYNTHESIS  
: NUMBER OF SEQUENCES: 9  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
: STREET: 1251 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: United States of America  
: ZIP: 10020  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/196,390  
: FILING DATE:  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 196 21 588.9  
: FILING DATE: 29-MAY-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 196 36 917.7  
: FILING DATE: 11-SEP-1996  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/EP97/02793  
: FILING DATE: 28-MAY-1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Haley, Jr., James F.  
: REGISTRATION NUMBER: 27,794  
: REFERENCE/DOCKET NUMBER: AGREVO-9  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212) 596-9000  
: TELEFAX: (212) 596-9090  
: INFORMATION FOR SEQ. ID NO.: 5:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 2825 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear

D	b	1828	- - - - - AGACCGCGCATCTGTCACAGCGCATCTGCACAACATGGAGTGGAAACCCCGAGG	1877
Q	y	833	GCAITGACATTGAGGAGTGGAACCCCACAGACCGACAAGTTCTCTGTCTGCCCTTACGACC	892
D	b	1879	TGACGCCCCACCTCAGTTCGGACGGCTACACCAACTTCTCTCTGAGAACCC- - - - -	1929
Q	y	893	AGAACAGCGTCTACGCGCGCAAGCGCGCGCAAGTAGGCGCTTCGACGCGCGAGCTGGGCC	952
D	b	1930	- - - - - TGNACTCCGCAAGCGCGCAGTGCAGAGAGCGCTTCGACGCGAGCTGGGCC	1980
Q	y	953	TGCTCTTGACCCCCAACCGCCCCCTTTTGCGCTTTCATCGGCGCGCTTCGAGAGACGAGAGG	1012
D	b	1981	TGCAGGTTCGCGCGCACGTGCGCTGCTCGGCTTCATCGGCGCGCTTCGAGCGCGCAGAAGG	2040
Q	y	1013	GTGTGACATCATCTTGGCGCGCTTCGCCACAGATCTCTGGCCACCCCAAGTTCAGATCG	1072
D	b	2041	GCCTTGAGATCATTCGCGGACGCCATGCCCTTGATCTGTG- --AGCAGACGTGCAGCTGG	2097
Q	y	1073	CCATCTCTGGGTACCGCAAGCGCGCTACAGAGAAGCTGGTGAAGCGCATCGGCACCAAGT	1132
D	b	2098	TGATGCTGGSCACCGGCGCCACGACCTTGGAGAGCATCTGCAGCACTTCGAGCGGGAGC	2157
Q	y	1133	ACAAAGGCGCGCAAGAGGGGTGGTCAAGTTCTCGGCGCGCTTGGCGACATGCTCACCG	1192
D	b	2158	ACCACGAACAAGTTCGCGGGTGGTGGGTTCTCGTGGCGCTTGGCGACCGGATCACGG	2217
Q	y	1193	CGSGCGCGCACTTCATGCTTGGTGGCTCGCTCGCGCTTCGAGCCCTTGGGGCTGATCCAGCTGC	1252
D	b	2218	CGGGGCGGACGCGCTCTCATGGCTCCCGGTTCGAGCGCTGGGGCTGAACCACTCT	2277
Q	y	1253	ACGCCATGCACTACGCTACCGTCCCGTGGTASCTTCACCGGGCGGCTGTGTGACACCG	1312
D	b	2278	ACGCCATGGCTTACGGCACGCTCCCGTCTGTCACAGCGCTCGGGGCTTCAGSGACACCG	2337
Q	y	1313	TCAAAGAGGCGTCAACGCTTCCAATCGGCGCGCTTGAACCCCGSACAAGCTGACACAGG	1372
D	b	2338	TGCGCGCGCTTCGACACCTTTCACCACTCCGCGCTCGGTGAGACGTTGACCGCCCGAGG	2397
Q	y	1373	CTGACGCGCGACGCGCTTGGCGCGCCACCGTTCGCGCGTTCGACGAGAGTGTTCGCGGCGGCC	1432
D	b	2398	CGCAACAAGCTGATCGAGGCGCTCGGSCACTGTGCTCCGCACTTACGAGACTTCAAGGAGA	2457
Q	y	1433	GCPACCCGAGATGTTGGCCAACTGCATCAGCCAGGACCTGTCTGTGTCRAAGCCCGGCC	1492
D	b	2458	GCTGGAGGGCCCTCCAGGAGCGGGCATGTCCGAGGACTTCAGCTTGGGAGCACGCCCA	2517
Q	y	1493	AGAAAGTGGGAGGCGCTGCTGGAGGAGGTGGTGTAC	1527
D	b	2518	AGCTCTACGAGGAGCTCTTCGTAAGGCCAAGTAC	2552
RESULT 8				
US-08-941-445A-10				
; Sequence 10, Application US/08941445A				
; Patent No. 6107060				
; GENERAL INFORMATION:				
; APPLICANT: Keeling, Peter				
; APPLICANT: Guan, Hanning				
; TITLE OF INVENTION: Starch Encapsulation				
; NUMBER OF SEQUENCES: 37				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.				
; STREET: 5370 Manhattan Circle				
; CITY: Boulder				
; STATE: CO				
; COUNTRY: US				
; ZIP: 80303				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: PatentIn Release #1.0, Version #1.30				
; CURRENT APPLICATION DATA:				

```

; APPLICATION NUMBER: US/08/941.445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2097 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2097
; US-08-941-445A-10

Query Match      11.1%; Score 216.8; DB 3; Length 2097;
Best Local Similarity 50.8%; Pred. No. 1.4e-29;
Matches 789; Conservative 0; Mismatches 657; Indels 108; Gaps 7;

QY 1 CGCGTGACATCGTGATGGTGTGCTGAGTGCSCCCTTGGTCCAGACGGGGCGGCTG 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 616 GTGATGAACGTGCTGCGTGGTCTGATGCTGCTCTTCTCAAGACAGGTGGCCT 675

QY 61 GCGGATGACTGGTGCCCTGCTATGAGCTAGTCAAGCGCGCCACCGCGTCATGACC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 GGAGATGCTGCTGGTGTCTTGGCTAAGGCTCTGCGAGGAGAGACACCGTGTATGGTC 735

QY 121 ATTGCCCTCGCTACGACCACTAGCTGAGCTGGGACACCTCGGTGGTCTGAGCATC 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 GTGATCAAGATGATGAGAGATGCGGAAGCCCGGG--ATTAGGTGTAAGGAGAGCTT 793

QY 181 ATGGGGGAGAAGTCCGCTACTTCCACTCCATCAAGAAGGCGCTGACCGGTGTGATT 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 ACAAGTAGCTGCACAGGATTCAGAGTACTTATTTTCACTCTTCAATGATGGAGTTG 853

QY 241 GACCACCCCTGGTTCTGGCCCAAGGTCTGGGGCAAGACCGGCTCCAAGCTGTACGGCCCC 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 ATTTTGTATTCTGTAGAAGCCCTCCCTTCGGGCAC-----CGSCACATAAT 900

QY 301 CGCTCCGGGCTGCTACTACCTGGACACCAACAGCGCTTCGCCCTGTCTCAAGGCGGCT 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 ATTTATGGGGGAAAGATGGATATTTTGAAGCGCATGATTTTGTCTCAAGGCGGCT 960

QY 361 ATTAGGCTGCCCGCTGCTGCCCTTCGGCC-----CCGCGAGGACTGGCTC 408
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 GTTAGGTTCCATGGTATGCTCCATGTGCGGCTACTGTCTATGGTGTATGCAACTTAGIT 1020

QY 409 TTCGGTGGCAAGACTGCGCACTCCGCCCTGGTGGCCGCTGCTGTGAAGACGAGTACCAG 468
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 TTCAATTGCTAATGATGGCATACCGCACTTCTGCCCTGTATCTAAAGGCTATTACCG 1080

QY 469 CCAAGGGCCAGTTCAACAGGCCAAGTGGTGGCTGGCTATCCACAACATGCGCTTCCAG 528
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 GACAATGGTTTATGATGAGTATGCTGCTGCTGTGATACACAACATGCTCATCAG 1140

QY 529 GSCCGCATGTGGGAGGAGCTTTCAAGGACACAGAACTGCCCGGAGCGCCCTTTGACAAG 588
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 GGTCTGGCCCTGTAGACACTTCGTCAATTTTGTACTTGCCTGAAACACTACATCGACCAC 1200

QY 589 CTGGCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGGCGCACCCCGCATGGAGGAGAC 648

```

## RESULT 9

```

US-08-572-951-3
; Sequence 3, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH

```

```

Db 1201 TTCAAACCTGTATGACAACAATTTGCTGGGATCAC----- 1233
QY 649 GAGAAAGCCCGCTGACGGGAAGACCTTACAAGAAGATCAACTGGCTGAAGGTGGCAATT 708
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1234 -----AGCAAGTTTTTCTCGGGGCTG 1257

QY 709 ATCGCCCGCGACAAGCTGTGTGACTGTGTCGCCCAACTACGCGACCGAGATGCTGCCGAT 768
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1258 AAGACGGCAGACCGGCTGTGACCGTTAGCATGCTACATGTGGAGCTGAAGACTTGG 1317
QY 769 GCCGCCGGGGTGTGGAGCTGGACACCGCTATPC-----GCGCAAGGGCAATTGAGGCAATT 825
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1318 GAAGCGGGTGGGCTTCCACGACATCAATAACAGAGACTGGAAGCTGCAGGGCATC 1377
QY 826 GTGAACGGCATGGACATTTGAGGAGTGAACCCCAAGACCGACAAGTTCTGTCTCGGCC 885
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1378 GTGAACGGCATGGACATGAGGAGTGAACCCCGCTGTGGAGCTGCACCTCCACTCCGAC 1437
QY 886 TACGACCAGAAC-----AGCGTCTACGCGCGCAAGCGCGCCCAAGAGAGGCC 933
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1438 GACTACACCAACTACACGTTTCGAGACGCTGGACACCGCGCAGCGCATGCAAGGCCGCC 1497
QY 934 CTGCAGCGCGAGCTGGGCTGCTGTGGAACCCACCGCCCGCTTTCGCCCTTCATCGGC 993
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1498 CTGCAGCGCGAGCTGGGCTGCTGTGCGCACCGGGCGCGACCTGATCGGGTTCATCGGG 1557
QY 994 CGCCTGGAGGAGCAGAGGGTGTGACATCATCTTGGGTACCGCAAGCGCGCTACGAAAGCTGTG 1113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1558 CGCTGGAACCAAGAGAGGGCTGACATCATCGCCGACGCGATCCACTGTGATC---GCG 1614
QY 1054 ACCCCCAGGTGCAAGTGCATCTTGGGTACCGCAAGCGCGCTACGAAAGCTGTG 1113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1615 GGGCAGGACGTGACGTGCTGTGCGCACCGGGCGGCGCGACCTGTGAGGACATGCTG 1674
QY 1114 AACGCCATCGSCACCAAGTACAAGGGCGCGCCCAAGGCGTGTGCAAGTTCTCGGCGGCC 1173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1675 CGGCGGTTCGAGTCGAGACACGAGCAAGGTGCGCGCTGGGTGGGGTTCTCGGTGCC 1734
QY 1174 CTGCGCACATGCTACCGCGCGCGCGACTTCATGCTGTGCGCTCGCGCTTCGAGGCC 1233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1735 CTGCGCACCGCATACGCGGGCGGCGGACATCTGCTGATGCGCTGCGGTTTCGAGCGC 1794
QY 1234 TCGCGCTGATTCAGCTGCACGCGCATGCTACGCTACCGTCCCGCTGGTGAAGCTTCCACC 1293
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1795 TCGGGCTGAACCAAGCTTACGCCATGGCGTACGGGACCGTGGCGGTGTCACGCGGTG 1854
QY 1294 GCGGCGCTGTCGACACCGTCAAGGAGGGCGTACCGGCTTCCACATGGGCGGCCCTGAAC 1353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1855 GGGGGGCTCGGACACGCTGGGCGCGCTTCAACCGCTTCAACGACACCGGCTCGGCTG 1914
QY 1354 CCGCACAAAGCTGGACGAGCTGACCGCGCGCGCTTGGCCCGCCACCGTGGCGGTGCCAGC 1413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1915 ACCTTCACCGCGCGAGGCGCAACCGCATGATCGACGCGCTCTGCACTGCCCTACCAACG 1974
QY 1414 GAGGTGTTTGGGGCGCGCTACCCCGAGATGTTGCCCAACTGCATCAGCGAGGACTG 1473
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1975 TACCGCAACTACAAGGAGAGCTGGCGCGCTGCGAGGCGCGCGCATGCGCGGAGGACTC 2034
QY 1474 TCCTGTGTCGAAGCCCGCCAGAGAGTGGAGGGCGCTGCTGGAGGAGGTGTGATC 1527
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2035 AGCTGGACACCGCCCGCTGCTGTATGAGGAGCTGCTGCTCAAGGCGCAAGTAC 2088

```









```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941.445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2007
; US-08-941-445A-8

Query Match      8.2%; Score 161; DB 3; Length 2007;
Best Local Similarity 53.2%; Pred. No. 6.4e-20;
Matches 444; Conservative 0; Mismatches 370; Indels 21; Gaps 4;

QY 699 GGTGTGCTATTCGCGCGGACAGCTGGTGGAGCTGGACACCGTC---ATCGCGCCCAAGGGCAT 815
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1170 GTGTGTTGAAGATGGACCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1229

QY 759 CGCTGCCGATGCGCGCGGGGTGGAGCTGGACACCGTC---ATCGCGCCCAAGGGCAT 815
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1230 GAAGACAGTGGAAAGCGGCTGGGCGCTCCACGACATCATCTGTTCAACACATGGAGAT 1289

QY 816 TGAGGGCATTTGAACGGCATGACATTTGAGGAGTGGAAACCCCAAGACGACAAAGTTCC 875
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1290 CAATGGCATTCGTGAACGATCGACCGATCGACACCAAGGAGTGGAAACCCCAAGTGGACCT 1349

QY 876 GTCTGCGCCCTACGACCAAGACAGCG-----TCTAGCGCGCAAGCGCCCGC 923
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1350 CGGTGGGACGGCTACACCACTACTCTCTCGAGACATCGACGCTGGAAAGCGGCAAGT 1409

QY 924 CAAGAGGCGCTGAGGCGGAGTGGGCGCTGCTGTGGACCCCAAGCGCCCGCTGTGCG 983
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1410 CAAGCGGCGCTGACGCGGAGCTGGGCGCTGGAAGTGGCGGAGCTGCGCGCTGCTCG 1469

QY 984 CTTTCATGCGCGCTGGAGGACGAGAGGTGTGGACATCATCTGCGCGCGCTGCCAA 1043
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1470 CTTTCATGCGCGCTGGATGGACAGAGGCGTGGACATCATCTGCGGAGCGGATGCGGTG 1529

QY 1044 GATCTGTGGCCACCCCAAGTGCAGATCCCATCTCTGGTACCGGCAAGCGCGCTACGA 1103
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1530 GATC---GGGGGACGAGGCTGACGCTGTGTGCTGGGACCGCGGCCACCTGACCTGGA 1586

QY 1104 GAAGCTGTGAAGCCATCGGCACCAAGTACAAAGGCGCGCCCAAGGGCGTGGTCAAGTT 1163
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1587 ACGAATGCTGCAGCACTTGGACCGGAGCATCCCAACAGGTGGCGGGTGGTGGTGGTGGT 1646

QY 1164 CTCGGGCGGCTGGGACATGCTCACGCGCGGCGCGGCTTCTGCTGCTGCTGCTGCTG 1223
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1647 CTCGGTCTTAATGATGATCGATCAAGCGCGGCGCGGCTGCTGCTGCTGCTGCTGCTG 1706

```

```

QY 1224 CTTGAGCCCTGCGGCTGATCCAGCTGACGCCATGCACTACGCTACGCTGCCGTGGT 1283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1707 CTTGCG---CCGGCGGCTGAACCACTCTACGCGATGCGATGCGATGCGTCCCTGTGGT 1763

QY 1284 AGCTTCCACCGCGCGCTGGTGGACACCGCTCAAGAGGCGGCTACCGGCTTCCACATGG 1343
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1764 GCACGCCCTGGCGGCTCAGGGACACCGTGGCGCGCTTGGACCGGCTTGGCGGCGG 1823

QY 1344 CGCCCTGAACCCGACAAAGCTGGACGAGCTACGCCACGCCCTGGCGCGGCGGCGG 1403
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1824 GCTCGGTTGGACTTTTACCGCGCGGAGGCAACAAAGCTGATCGAGGTGCTCAGCACTG 1883

QY 1404 CGTGCACGAGGTGTTTGGCGGCGCGCTACCCCGAGATGGTGCCCACTGCATCAG 1463
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1884 CCTGACACGTTACCGAACTACGAGGAGAGTGGAGAGTCTCCAGGCGCGGCGGCTGTC 1943

QY 1464 CCAGGACTGCTGCTCCAAAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 1518
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1944 GCAGAACCTCAGCTGGGACCGCGGCTGAGCTCTACGAGGAGCGTCTTGTCAAG 1998

```

## RESULT 14

```

US-08-572-951-2
; Sequence 2, Application US/08572951
; Patent No. 5824790
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L.
; APPLICANT: KNIGHT, MARY E.
; APPLICANT: GUAN, HANPING
; TITLE OF INVENTION: MODIFICATION OF STARCH
; TITLE OF INVENTION: SYNTHESIS IN PLANTS
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: Pillsbury Madison & Sutro LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/572,951
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,602
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/263,921
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul N. Kokulis
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 222957/1.02.15C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

US-08-572-951-2

Query Match 8.2%; Score 161; DB 1; Length 2085;  
Best Local Similarity 53.2%; Pred. No. 6.4e-20;  
Matches 444; Conservative 0; Mismatches 370; Indels 21; Gaps 4;

QY 699 GGGTGGCTATATCGCGCGGCAAGAGTGGTACTGTGTCGCCAACTACGCGACCGAGAT 758  
DB 985 GTGTGTTCTGAAGATGGACAGCGGTGTGACTGTACGCGGGCTACCTGTGGAGCT 1044

QY 759 CGTGGCCGATGCGCGCGGGTGTGAGCTGGACACCGTC---ATCGCGCCCAAGGGCAT 815  
DB 1045 GAAGACAGTGAAGCGGGCTGGGGCTCCACGACATCATCCGTTTAAAGACTTGAAGAT 1104

QY 816 TGAGGCGATGTGAACGGGATGACATGAGGAGTGAACCCCAACACGACGACAAGTTCT 875  
DB 1105 CAATGGCATTTGTAACGATGACACCAAGAGTGAACCCCAAGAGTGAACGTCACCT 1164

QY 876 GTCTGCGCCCTAGACCAAGACAGCG-----TCTAGCGCGGCAAGCGCCCGC 923  
DB 1165 GCGGTGCGAGGCTACACCAACTACTCTCTCGAGACACTCGAGCTTGAAGCGGAGTG 1224

QY 924 CAAGGAGCGCTCGACGCGGAGTGGGCTGCTGTGGACCCACCGCCCTGTTCGC 983  
DB 1225 CAAGGCGGCGCTCGACGCGGAGTGGGCTTGAAGTGGCGAGCGCTGCGCTGCTCG 1284

QY 984 CTTTCATCGCGCGCTGAGGAGCAGAGGTTGGACATCATCTGGCCGCTGCGCCAA 1043  
DB 1285 CTTTCATCGCGCGCTGAGGAGCAGAGGCTGGACATCATCTGGGACCGGATGCGGTG 1344

QY 1044 GATCTGGCCACCCCAAGTGCAGATTCGCCATCTCTGGTACCGCAAGCGCGCTACGA 1103  
DB 1345 GATC---GCGGGGACGAGCTGAGCTGTGTGCTGGGACCGCCACCTGACCTGGA 1401

QY 1104 GAGCTGGTGAAGCCATCGGACCAAGTACAGAGGCGCGCAAGGGCGTGTCAAGTT 1163  
DB 1402 ACGAATGCTGCAGCACTTGGAGCGGGAGCATCCCAAGGTCGCGGGTGGTGGTGGTT 1461

QY 1164 CTGGGCGCCCTGGCGACATGCTACCGCGCGCGGAGTCTATGCTGGTGGCTCGCG 1223  
DB 1462 CTGGGCTTAATGTTGATCGCATCAGCGCGGGCGGCGAGTGTGTGATGCGCTCCCG 1521

QY 1224 CTTGAGCGCTGGCGCTGATCGAGTGCAGCGCATGACATGACGTTACCGTGGCCGCTGGT 1283  
DB 1522 CTTGCG---CGGGGCGGTGAACAGCTCTACGGGATGGCATACGCGACCGCTCCCTGGT 1578

QY 1284 AGCTCCACGCGCGCTGTGTCACACCTCAAGGAGGCGTCAACCGCTTCCACATGG 1343  
DB 1579 GCAGCGGTGGCGGGCTCAGGACACACCGTGGCGCGCTTCCGCGACCGCGG 1638

QY 1344 CGCCCTGAACCCGACAGCTGGAGGAGTGGAGCGCGGCTGGCGCGGCGGCTGGCG 1403  
DB 1639 GCTGGGTGACATTTTACCGCGCGGAGGCGCAACAGCTGATCGAGTGTCTACCGACTG 1698

QY 1404 CGGTGCGAGCGAGTGTGTCGGGCGCGCGCTACCCGAGATGTTGGCCAACTGCATCAG 1463  
DB 1699 CCTGCACACGTACCGAAACTACGAGGAGAGTGAAGAGTCTCCAGGCGCGCGCATGTC 1758

QY 1464 CCAGGACTGCTCTGTTCCAGACCGCGGCAAGTGGAGGCGCTGCTGGAGGAG 1518  
DB 1759 GCAGAACCTCAGCTGGGACCAACCGGCTGAGCTCTACGAGGAGCTCTTGTCAAG 1813

RESULT 15  
US-09-196-390-1  
; Sequence 1, Application US/09196390  
; Patent No. 6307125  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; APPLICANT: Lorz, Horst  
; APPLICANT: Lutticke, Stephanie  
; APPLICANT: Walter, Lennart  
; APPLICANT: Froberg, Claus

APPLICANT: Kossmann, Jens  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH  
TITLE OF INVENTION: SYNTHESIS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,390  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 21 588.9  
FILING DATE: 29-MAY-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 36 917.7  
FILING DATE: 11-SEP-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP97/02793  
FILING DATE: 28-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2239 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORGANISM: Triticum aestivum L.  
STRAIN: cv. Florida  
HAPLOTYPE: ca. 21 d Caryopses  
IMMEDIATE SOURCE:  
LIBRARY: cDNA library in pBluescript sk (-)  
CLONE: Tasss  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..2017  
US-09-196-390-1

Query Match 6.1%; Score 118.2; DB 4; Length 2239;  
Best Local Similarity 52.4%; Pred. No. 1.6e-12;  
Matches 333; Conservative 0; Mismatches 293; Indels 9; Gaps 3;

QY 686 TCAACTGGCTGAAGGTGGCATATTCGCGCGGAGTGGAGTGGAGTGGACACCGTCATCC 745  
DB 823 TTAACCTTTTGAAGGAGCAGTGTGACAGCAGATCGGATTTGACCGCTCAGTACAGGTT 882

QY 746 ACGGACCGGAGATC---GCTGCGGATCGCGCGGGTGTGGAGTGGACACCGTCATCC 802  
DB 883 ATTCAATGGAGGTCAACACTGCTGAAGTGGACAGGCGCTCATGAGCTCTTAAGCTCC 942

QY 803 GCGCCAAAGGCGATTGAGGCGCAATTGTGAACGGCATGACATGAGGAGTGAACCCCAAGA 862  
DB 943 GAAAGAGTGTATTGAATGAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGA 1002



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 21:49:59 ; Search time 177.291 Seconds  
(without alignments)  
15416.986 Million cell updates/sec

Title: US-09-980-771A-4

Perfect score: 1953

Sequence: 1 gcgtgacatcgatggt.....ccgccaccgcgaagtcgcgc 1953

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	475.8	24.4	2267	9	US-09-961-077-25
2	314.6	16.1	2399	9	US-10-138-075-1
3	275.2	14.1	2179	9	US-10-138-075-3
4	218.2	11.2	2825	10	US-09-952-677-5
5	118.2	6.1	2239	10	US-09-952-677-1
6	102	5.2	2712	10	US-09-748-033-4
7	94.2	4.8	790	10	US-09-966-881-44
8	86.6	4.4	1266	9	US-10-145-415-98
9	84.2	4.3	2307	9	US-09-893-519A-87
10	83.2	4.3	1427	10	US-09-974-300-757
11	82.2	4.2	805	10	US-09-966-881-43
C 12	81	4.1	12441	9	US-09-988-384B-3
C 13	81	4.1	13613	9	US-09-860-846-3
C 14	81	4.1	13613	9	US-09-836-821-3
C 15	81	4.1	13613	10	US-09-861-289-3
C 16	76.4	3.9	813	10	US-09-966-881-45
C 17	76	3.9	1318	9	US-09-934-900-15
C 18	75.4	3.9	446	10	US-09-864-761-20699
19	75.2	3.9	1896	9	US-10-124-880-15

20	74.8	3.8	1407	10	US-09-815-242-4089	Sequence 4089, Ap
21	74.8	3.8	15872	9	US-09-860-846-1	Sequence 1, Appli
22	74.8	3.8	15872	9	US-09-988-384B-1	Sequence 1, Appli
23	74.8	3.8	15872	9	US-09-836-821-1	Sequence 1, Appli
24	74.8	3.8	15872	10	US-09-861-289-1	Sequence 1, Appli
25	74.6	3.8	1266	9	US-10-145-415-104	Sequence 104, App
26	74.2	3.8	1770	9	US-10-100-957A-1	Sequence 1, Appli
27	73.6	3.8	88421	9	US-09-976-059-1	Sequence 1, Appli
28	72.4	3.7	3468	9	US-09-988-462-2	Sequence 2, Appli
29	72.2	3.7	299	10	US-09-294-093B-3520	Sequence 3520, Ap
30	72.2	3.7	824	10	US-09-894-633A-83	Sequence 83, Appli
31	71.4	3.7	3624	9	US-09-988-462-6	Sequence 6, Appli
C 32	71.4	3.7	12733	9	US-10-032-393-47	Sequence 47, Appli
C 33	71.4	3.7	12739	9	US-10-032-393-8	Sequence 8, Appli
34	71.2	3.6	1182	9	US-09-364-847-1	Sequence 1, Appli
35	71.2	3.6	1926	9	US-09-364-847-9	Sequence 9, Appli
36	71.2	3.6	1926	9	US-09-364-847-18	Sequence 18, Appli
37	70.8	3.6	1415	9	US-09-934-900-11	Sequence 11, Appli
38	70.6	3.6	1635	10	US-09-864-761-20241	Sequence 20241, A
39	70.6	3.6	1973	10	US-09-864-761-3471	Sequence 3471, Ap
C 40	69.8	3.6	1075	10	US-09-864-761-19241	Sequence 19241, A
C 41	69.8	3.6	1403	10	US-09-864-761-2513	Sequence 2513, Ap
42	69.8	3.6	2010	12	US-10-032-717-9	Sequence 9, Appli
43	69.8	3.6	4689	9	US-09-860-846-34	Sequence 34, Appli
44	69.8	3.6	4689	9	US-09-988-384B-34	Sequence 34, Appli
45	69.8	3.6	4689	9	US-09-836-821-34	Sequence 34, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-961-077-25  
; Sequence 25, Application US/09961077  
; Publication No. US20030014775A1

#### GENERAL INFORMATION:

APPLICANT: Zwick, Michael G.  
Edington, Brent E.  
McSwiggen, James A.  
Merlo, Patricia Ann Owens  
Guo, Lining  
Skokut, Thomas A.  
Young, Scott A.  
Folkerts, Otto  
Merlo, Donald J.

TITLE OF INVENTION: COMPOSITION AND METHODS FOR  
MODULATION OF GENE EXPRESSION  
IN PLANTS

NUMBER OF SEQUENCES: 1263

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/961,077

FILING DATE: 21-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/679,645

FILING DATE: July 12, 1996

APPLICATION NUMBER: 60/001,135

FILING DATE: July 13, 1995

APPLICATION NUMBER: 08/300,726

FILING DATE: September 2, 1994

## ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 219/247

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

## INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 2267 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-961-077-25

Query Match 24.4%; Score 475.8; DB 9; Length 2267;  
Best Local Similarity 59.9%; Pred. No. 5.7e-99;  
Matches 979; Conservative 0; Mismatches 557; Indels 99; Gaps 7;

QY 5 TGGACATCGTATGTTGCTGAGGTGCGCCCTTGGTCCAGAGCGGCGCTGGCG 64  
DB 418 TGAACGCTCTCTGTCGGCGCGAGATGCGCCCTGGAGCAAGACCGCGCTCGCG 477

QY 65 ATGTGACTGTGCTGCTATTTAGCTGCTCAAGCGCGCCACCGCGTCATGACCATG 124  
DB 478 AGCTCTCTGCGGCGCTTGGCGCGCGCATGCGCGCAATGGGCACTGTCATGCTGTCT 537

QY 125 CCCTCGCTACGACCACTAGCTGACGCTTGGGACACCTCGTGGTGGACAT----- 179  
DB 538 CTCGCCCTAGGACCACTCAAGGACGCTGGGACACCGCGCTGTCGAGATCAAGA 597

QY 180 -----CATGGCGAGAGTCCGCTACTTCCATCCATCAAGAGGGCGTCCACGCG 232  
DB 598 TGGGAGACAGGTACGAGCGGTCAAGTTCTTCCACTGTACAAAGCGCGAGTGGACGCG 657

QY 233 TGTGGATTGACCACTGCTTCTTGGCAAGGTCTGGGCAAGACCGGCTCCAAAGTGT 292  
DB 658 TGTCTGTGACCACTGCTTCTGAGAGGTTTGGGAAAGACCGGAGAGAGATCT 717

QY 293 ACGGCGCGCTCCGCGCTGACTACTGTGACAAACCAACAGCGCTTGGCCCTGTCTGCA 352  
DB 718 ACGGCGCTGACGCTGGAACGACTACAGGACAAACAGCTGCGGTTGAGCTGCTATGCC 777

QY 353 AGCGCGCTATTAGGCTGCGCGCTGCTGCCCTTGGC----- 390  
DB 778 AGGACGACTTGAAGTCCAAAGTCTTGAAGCTTCAACAAACCCCAATCTTCTCGGAC 837

QY 391 --CCCGCGAGGACTGCTTCTGTCGCAACGACTGGCACTCGCCCTGTCGCGCTCC 448  
DB 838 CATACGGGAGGAGCTGCTTCTGTCGCAACGACTGGCACTCGCCCTCTCTCTGTCT 897

QY 449 TGTGTAAGGACGAGTACCAAGGCGGAGGCGGCTTCCAGAGGCGGCTGCTGCTGCTA 508  
DB 898 ACCTCAAGACCACTACCACTCCCGGCACTACAGGAGCGCAAGACCGCTTCTGCA 957

QY 509 TCCACACATCTGCTTCCAGGCGCGATGTTGGAGGAGGCTTTCAGGACACAGAGCTGC 568  
DB 958 TCCACACATCTCTTACCAGGCGGCTTCCGCTTCTCCGACTACCGGAGCTGAACCTCC 1017

QY 569 CCCAGGCGGCTTTGACAAAGTGGCTTCTCGGACGCTATGCAAGGTTTACACTGAGG 628  
DB 1018 CGGAGAGATCAAGTCTGCTTCTGATTCATCGACGCTACGAGAAG----- 1064

QY 629 CCACCCCATGGAGGAGGACGAGAGCCCGCTGACGGGAAAGACCTACAGAGATCA 688  
DB 1065 -----CCCGTGGAG-----GCGGGAAGATCA 1086

QY 689 ACTGGCTGAAGGTTGGCATTTATCGCGCGGACAGCTGTTGACTGTGTCGCCCACTACG 748  
DB 1087 ACTGGATGAAGCGCGGATCTCGAGCGGACAGGGTCTTCCACGTCAGCCCTTACTACG 1146

## RESULT 2

US-10-138-075-1

; Sequence 1, Application US/10138075

; Publication No. US20030087369A1

; GENERAL INFORMATION:

; APPLICANT: Broglie, Karen E.

; APPLICANT: Butler, Karlene H.

; APPLICANT: Harvell, Leslie T.

; APPLICANT: Lightner, Jonathan E.

; APPLICANT: Orozco, Emil M.

; TITLE OF INVENTION: Granule-Bound Starch Synthase





```

; TYPE: DNA
; ORGANISM: Glycine max
US-10-138-075-3

```

Query Match 14.1%; Score 275.2; DB 9; Length 2179;  
Best Local Similarity 52.7%; Pred. No. 1.6e-53;  
Matches 820; Conservative 0; Mismatches 628; Indels 108; Gaps 6;

11	TCGTGATGTTGCTGCTGAGTTCGCCCTTGGTCCAAAGACGGCGCCCTGGCGGATGGA	70
Db		
323	TCATAAATATAGGAACCGAGTGGCTCCCATGTTGCAAAACTGGTGGTGGGAGATGTC	382
Db		
71	CTGGTGGCCTGCTATTAGCTGGTCAAGCGGGCCACCGCGTCATGACCATTTGCCCTC	130
Db		
383	TTGGAGGTCATCCACCGGCATTGGCAGGTTTTGGGCATCGAGTAATGACTATTGTGCCGC	442
Db		
131	GCTACGACCACTACGCTGACGCTGGGACCACTCGGTGGTCTGTGACATCATGG-----	184
Db		
443	GCTATGACCACTACAAAGATGATGGGATACAAAGTGTGTAAATTGAGGTGAAAGTAGGAG	502
Db		
185	-----GCGAGAAAGTCCCTACTTCCACTCCATCAAGAAGGGCGTGCAACCGCTGTGGA	238
Db		
503	ATAGAACAAGAAAGTTCCCTTCTTCCATTGTTATAGAGGGAGTTGATCGTGTCTTTG	562
Db		
239	TTGACCAACCCCTGGTTCTCGGCCAAGTCTGGGGCAAGACCGGCTCCAAGCTGTACGGCC	298
Db		
563	TGGATCACCCCTTGGTTCTTGAAGAGGTATGGGGAAAACTGGACAAAAACTTATGGAC	622
Db		
299	CCGCTCCGGCGCTGACTACTCTGGACAACCAACGAGCGCTTCGCCCTGTCTCGAAGCCG	358
Db		
623	CAACTACTGGAAATGATTAGAAAGACAACCAACTGCGTTTACGCTCTTTGCCAGGCTG	682
Db		
359	CTATTGAGGCTGCCCGTGTCTCCCTTCGGCCCC-----G	394
Db		
683	CTTTGGAAGCCCCAAGGTTCTGAGTCTTAATTCAGTAAATATTTCTCGGACCATATG	742
Db		
395	CGGAGGACTCGCTCTTCGTGGCCCAAGACTGGCACTCCGCGCTGTGGCCGTCCTCTGA	454
Db		
743	GTGAAGATGCAATTTTGTGGCAATGATTGGCACACTGCCCTTATCCCTGCTCACTTGA	802
Db		
455	AGGACGAGTACCAAGCCCAAGGGCCAGTTACCAAGGCCAAGTCGGTGTGGCTATCCACA	514
Db		
803	AAAGTATGTACCAAGTCAAGGGGATCTATACGAATGCCGGGTGTTTTTGTATATCCACA	862
Db		
515	ACATCGCTTCCAGGGCCGATCTGGGAGGAGGCTTTCGAAGGACAGGAAGTCCGCCAGG	574
Db		
863	ACATTGCTTACCAAGGAGATTTGCATTCCCGACATCTCACITCTTAATCTCCAGACC	922
Db		
575	CCGCTTTGACAGCTGGCCTTCTCGGACCGGCTATGCCAAGGTTTACACTGAGGCCAACC	634
Db		
923	AAITTAAGAGCTCCTTTGACTTTATTGATGGGCATG-----	958
Db		
635	CCATGGAGGAGGACGAAGACCCCGCTGACGGGAAGACCTACAGAAGATCAACTGGC	694
Db		
959	-----TTAAACACAGTGGTTGGAAGAAATCAATTGGT	991
Db		
695	TGAAGGTGGCATTATCCGCCCGCACAAAGCTGGTGTGTGTCGCCCAACTAGCGACCG	754
Db		
992	TGAAGCTGGACTTATAGATCATGTTTGTGATAACCGTTAGTCCAAACTATGCTAAAG	1051
Db		
755	AGATCGCTGCCGATGCCCGCGGGGTGTGAGCTGGACACCGCTATCCCGCGGCAAG---	811
Db		
1052	AAC'TGGTGTCAAGTCCAGACAAGAGGATGGAATTTGGACAACATCATTCGCAAAATGATG	1111
Db		
812	-----GCATTTGAGGGCATCTGAACGGCATGGACATTGAGGAGTGGAAACCCCAAGA	862
Db		
1112	ATGATGTCGTTTGTGTTGGAATTTGGAATGGCATGGATGTTTCAGGAGTGGAAATCCACCA	1171
Db		
863	CCGACAAGTTCTCTGTGCGCCCTTACCAACAGAACACGCTCTACCGCGGCAAGCCCGCG	922
Db		
1172	CTGACAAATATATAGCTGTCAATATATGATGTTTCAACAGTATTCGAAGCAAGGCTCTTC	1231
Db		
923	CAAAGGAGCCCTCGACGCGAGCTGGGCTGCTGTGGACCCCAACCGCCCTGTGTTCG	982
Db		

Db	1232	TGAAGAAGCCCTCCAAAGCAGAAGTTGGATTCCCACTCGACAGAAATATTCCTCTCATTG	1291
Qy	983	CCATTATCGCCCGCTGGAGGAGCAGAAGGTTGGACATCATCCTGGCCGCCCTGCCCA	1042
Db	1292	GTATTCAITGTGTAGCTTGAAGAGCAAAAGGTTCTGAATATCTTGCAGAAGCCATCCCC	1351
Qy	1043	AGATCTCGGCCACCCCAAGGTGCAGATCGCCATCCTGTGGTAGCGGCAAGCGCGCCTACG	1102
Db	1352	AATTTAT--CAAGCAGAAATGTTCAAGTTGGTAGCCCTAGCAAGGAAAAACAATGG	1408
Qy	1103	AGAAGCTGTTGAACGCCATCGGCACCAAGTACAAAGGCCGCCGCAAGGGCGTGGTCAAGT	1162
Db	1409	AAAAGCAGCTTGAGGAACTTGAAATATCATACCTGATAAGGCCAGAGGAGTGGCAAAAT	1468
Qy	1163	TCTCGGGCCCTTGGCCACATGCTCACCGCCGGCCGCGACTCATGCTGCTGCGCCCTCGC	1222
Db	1469	TCAATGTTCCCTAGCCCAATGATAATAGCTGGAGCTGATTTATATTGTTTCCCTAGCA	1528
Qy	1223	GCATCGAGCCCTGGCGGCTGATCCAGCTGCACGCCATCGCATAGSGTACCGTGCCTGG	1282
Db	1529	GATTTGAGCCTTGGGTCTCATTCAGTTCAAGCTATCGCTATGGATCTGTACCAATTG	1588
Qy	1283	TAGCCTCCACCGGGCGCTGTGCACACCGTCAAGGAGGGGCTACCGGCTTCCACATGG	1342
Db	1589	TTGCCCTCAACAGGTGGATTAGTTGACACTGTCAAAGAAGGCTTCACTGGATTTCAAGTGG	1648
Qy	1343	CGCGCCCTGA-----ACCCCGACAGCTGGAGAGGCTGACGCGCGAGCGCCTGGCCGCCA	1396
Db	1649	GTGCTTCAATGTTGAATGATGCTGTGGATTCGGCTGATGTGGATGCTATATCAAGA	1708
Qy	1397	CCGTGCGCGGTGCCAGCGAGGTGTTTGGCGGCGCGCGTACCCCGAGATGGTGGCCCAAT	1456
Db	1709	CTGTCAAAAGGCCCTTGCAGTCTATGGAATCCAGCTTTTACAGAAATTTACAAGAACT	1768
Qy	1457	GCATCAGCCAGGACCTGTCTGTGTCGAAGCCGCCCAAGAGTGGAGGGCGCTGCTG	1512
Db	1769	GCATGGCTCAAGATCTTTCATGGAAGGGCGCTGTGTAAGGATGGGAGGAAGTGTG	1824

RESULTS: A

```

RESOL.
US-09-952-677-5
; Sequence 5, Application US/09952677
; Patent No. US20020138876A1
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; Loiz, Horst
; Luticke, Stephan
; Walter, Lennart
; Frohberg, Claus
; Kossmann, Jens
;

```

**TITLE OF INVENTION:** NUCLEIC ACID MOLECULES ENCODING ENZYMES FROM WHEAT WHICH ARE INVOLVED IN STARCH SYNTHESIS

NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

RESPONSE ADDRESS: ADDRESS: James F. Haley, Jr., c/o Fish & Neave

ADDRESSEE: JAMES F. HALEY, JR., C/O  
STREET: 1251 Avenue of the Americas

STREET: 1231 AV  
CITY: New York

CITY: NEW YORK  
STATE: NEW YORK

STATE: NEW YORK  
COUNTRY: United States of America

COUNTRY: U  
ZTP: 10020

ZIP: 10020  
 UTTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
MEDIUM LIFE: floppy disk

COMPUTER, IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0

SOFTWARE: FALCONET  
 TEST APPLICATION DATA:

APPLICATION NUMBER: US/09/952,677

FILING DATE: 14-Sep-2011

PER APPLICATION DATA:

APPLICATION NUMBER: 09/196,390

FILING DATE: 19-NO. US20020138876A1-1998

```

1  APPLICATION NUMBER: DE 196 21 588.9
2  FILING DATE: 29-MAY-1996
3  APPLICATION NUMBER: DE 196 36 917.7
4  FILING DATE: 11-SEP-1996
5  APPLICATION NUMBER: PCt/EP97/02793
6  FILING DATE: 28-MAY-1997
7  ATTORNEY/AGENT INFORMATION:
8  NAME: Haley, Jt., James F.
9  REGISTRATION NUMBER: 27,794
10 REFERENCE/DOCKET NUMBER: AGREVO-9
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: (212) 596-9000
13 TELEFAX: (212) 596-9090
14 INFORMATION FOR SEQ ID NO: 5:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH: 2825 base pairs
17     TYPE: nucleic acid
18     STRANDEDNESS: single
19     TOPOLOGY: linear
20     MOLECULE TYPE: cDNA to mRNA
21     HYPOTHETICAL: NO
22     ANTI-SENSE: NO
23     ORIGINAL SOURCE:
24     ORGANISM: Triticum aestivum L.
25     STRAIN: cv. Florida
26     TISSUE TYPE: ca. 21 d Caryopses
27     IMMEDIATE SOURCE:
28     LIBRARY: cDNA library in pBluescript
29     CLONE: pTAS1
30     FEATURE:
31     NAME/KEY: CDS
32     LOCATION: 162..2559
33     SEQUENCE DESCRIPTION: SEQ ID NO: 5:
34     US-09-952-677-5

```

Query Match	11.2%;	Score 218.2;	DB 10;	Length 2825;
Best Local Similarity	50.1%;	Pred. No. 1.3e-40;		
Matches 769;	Conservative	0;	Mismatches 688;	Indels 78; Gaps 6;
QY	5	TGACATCGTGATGTTGCTGCTGAGTTCGCCCTTGCTCCAAAGCGGGCGCTGGGCG	64	
Db	1084	TGAACGTGGTCGTGGTCTGAGTGTCTCCCTGGTCAAAACAGGTGGCTGGGAG	1143	
QY	65	ATGTGACTGGTGCCTGCCCTATTGAGTGTGTCAAGCGCGCCACCGGTCATGACCATTTG	124	
Db	1144	ATGTTGGGCTGCTTGCCCAAGGCTTTGGCAAGAGAGACATCGTGTATTGTTGTGG	1203	
QY	125	CCCTCGCTACGACCACTAGCTGAGCGCTGGGACACCTCGCGTGGTCTGGACATCATGG	184	
Db	1204	TACCAAGGTATGGGACTATGAAGAAGCTTACGATTCGGAGTCCGAAATACTAACAGG	1263	
QY	185	GGCAGAAGTCCGCTACTTCCACTCCATCAAGAAGGCGGTGCACCGCGTGTGGATTGACC	244	
Db	1264	CTG-----CTGGACAGGATATGGAAGTGAATTAFTTCCATGCTTAATCGA	1309	
QY	245	ACCCCTGGTTCCTTGGCCAAAGGCTTGGGGCAAGACCGCTCCAAAGCTGTAAGGCCCGCGT	304	
Db	1310	TGGAGTGTATTTGTGTTTCATTGACGCTCCTCTCTTCCGAC-ACCCTCAGGAAGACATTT	1368	
QY	305	CCGGCGCTGACTACCTGGACACACCAAGCGTTCCGCTGTCTGTCTCAAGGCGCTATTG	364	
Db	1369	ATTGGGGCAGCAGACAGGAATATGACGCGCATGATTTGTCTGCAAGGCGCTGTG	1428	
QY	365	AGGTCGCCGCGTGTGCCCTTTCGGCCCGCGGAGACTGCG-----TCTTCG	412	
Db	1429	AGGTTCCATGGCAGCTTCCATCGCGGGTGTCCCTATATGGGATGGAATCTGGTGTATA	1488	
QY	413	TGGCCCAACGACTGGCACTCCGCCCTGGTGGCCGTCGTCTGAAGGACGAGTACCAAGCCCA	472	
Db	1489	TTGCAAAATGATTGGCAGACGGGACCTCCCTGTCTATCTGAAGCATATTACAGGGACC	1548	
QY	473	AGGGCCAGTTCACCAAGGCCAAGTCGGTCTGGCTATCCCAACATPGCCTTCCAGGGCC	532	

D	b	1549	A T G G T T G A T C G A G T A C A C T C G G T C C A T T A T G T G T A T A C A T A P A C A T C G C T C A C C A G G G C C	1608
Q	y	533	G C A T T G G G A G G A G C T T T C A A G A C A C A G A G C T G C C C A G G C G C C T T T C A C A A G C T G G	592
D	b	1609	G T G G C C C N T G A T A C A A T T C C G T T C A C C G A G T T G C C T G A G C A C T A C C T T G G A C A C T T C A	1668
Q	y	593	C C T T C T C G G A G G C G T A T G C C A A G G T T T A C A C T G A G G C C A C C C C A C T G A G A G G A G A C G A G A	652
D	b	1669	G A C T G T A G A C A C C C G T G G T G G T G A A C A C A C G C A A C T A C T T C G C C G C G G C C T G A A G A T G G	1728
Q	y	653	A G C C C C G C T C A C G S G A A G A C C T T A C A G A R A G A T C A A C T G C T G A A G G T G G C A T T A T C G	712
D	b	1729	C G G A C C A G T T G T C G T G T G A G C C C G G G T A C C T T G T G G A G A C T G A A C A C G G T G G A G G C G	1788
Q	y	713	C G C C G C A A A G C T G T G A C T G T F C G C C C A A C T A C G G A C C G A G A T C G C T G C G A T C C G	772
D	b	1789	G C T G G G G C T T C A G A C A T A T A C G G C A G A C G A C T G A	1827
Q	y	773	C C G G G G T G T G A C T G G A C A C G T C A T C G G C C A A G G C A I T G A G G C A T T G T G A A G C	832
D	b	1828	-----A G A C C C G G G A T C G T C A C G G C A T C G A C A A C A T G G A G T G A A C C C G A G G	1878
Q	y	833	G C A T G G A C A T T A G G A C T G G A A C C C A A G A C C G A A G T T C C T G T C G G C C C T A C G A C C	892
D	b	1879	T G G A G C C C A C C T C A A G T C G G A C G G T A C A C A A C T T C C C T G A G A C G C	1929
Q	y	893	A G A C A C G C T T A G C C G G A A G C G C C C C A A G A G G C C C T G C A G G C G A G C T G G G C C	952
D	b	1930	-----T G G A C T C C G G A A G C G G C A G T G C A A G A G G C C C T G C A G C G C G A G C T G G C C	1980
Q	y	953	T G C T G T G G A C C C A C C G C C C C T G T G C C T T A T C G C C T G A T C G G C G C C T G A G A G A C A G A A G G	1012
D	b	1981	T G C A G T C C C G C G A G T C C C C T G T G G C T T C A T C G G C G C C T G S A G G G C C A A G G	2040
Q	y	1013	G T G T G G A C A T A C T T G G C C C C C C C A A G A T C C T G G C C A C C C C A A G T G C A G A T C G	1072
D	b	2041	G C G T G G A G A T A T C G C G G A C G C A T C C C T G G A T C G T - - A G C C A G A C G T G C A S C T G G	2097
Q	y	1073	C C A T C T T G G T A C C G C A A G G C C C T A C A G A A G C T G T G A A C C C A T C P G G C A C C A A G T	1132
D	b	2098	T G A T C T G G C A C C G G G C G C A G A C T G G A G A C A T G C T G A G A C A T T C G A S C G G G A G C	2157
Q	y	1133	A C A A G G C C G C G C C A A G G G C T G G T C A A G T T C T G G C G C C C T G C G C A C A T G C T C A C G	1192
D	b	2158	A C C A C A A A G T G C G C G G T G G T G G G T T C T C G T G C C T G S C C A C C G G A T C A C G G	2217
Q	y	1193	C C G C C C G A C T A C T G T G T G C C T C G C G T T C A G A C C C T G C G G C C T G A T C C A G C T G C	1252
D	b	2218	C G G G G C G A C G C G C C T A T G C C C T C C C G T T G A G C C G T G C G G G C T G A A C A G C T C T	2277
Q	y	1253	A C C C A T G A C T A C G T A C C G T G C C G T G T A G C C T C C A C G C G G C C C T G T G C A C A C G	1312
D	b	2278	A C G C C A T G C C T A G C A C C G T C C C G T G T C A C C C T C C C G T T G A G C C G T C A G G A C A C G	2337
Q	y	1313	T C A A G A G G G C G T C A C C G C C T T C A C A T G G G G C C C T G A A C C C C A A A G C T G A C A G G	1372
D	b	2338	T G C C G C G T T C A C C C C T T C A A C C A C T C C G G C T C G G T G A C G T T C A C C G C C C A G G	2397
Q	y	1373	C T G A C C C C A C C C T G G C G C C A C C G T G C C G T G C C A G A G G T G T T G C G G G C G C C	1432
D	b	2398	C G C A C A A G C T A T C A G G C G C T C G G G C A C T G C C T C C G C A C C T A C C G A G A T T C A A G A G A	2457
Q	y	1433	G T A C C C G A G A T G T G C C A A C T C A T C A C C A G G A C T C T C C T G T C C A A G C C C C C C	1492
D	b	2458	G C T G A G G G C C C T C C A G A G G C G S C A T G T C G A G G A C T T C A G C T G G A G A C A C C C C C A	2517
Q	y	1493	A G A A G T G G A G G C C T G C T G A G A G G T G T G T A C	1527
D	b	2518	A G C T A C A G A A G C T C T C T C A A G G C C A A G T A C	2552

RESULT 5  
US-09-952-677-1



Db 975 GCAGCGCTTCGAGGCAAGTGAGCGTGAAGACACCGGACCGCCCGCCCTCAGGGCT 1034  
QY 689 ACTGGCTGAGGGTGGGATATCGCCCGGACAGCTGGTGAATGTGCGCCCAACTACG 748  
Db 1035 GGACCTCTGGAGTGGGACTTCGCCCGCGGAACAAAGGTGACCTCGGCTGGGACCGGACG 1094  
QY 749 CGACCGAGATCGCTGCGGATGCGCGCGGCTGTGGAGCTGGACACCGTCAATCGCGGCA 808  
Db 1095 TCACCAACAGCGGACCACTGGACCGCAAGAACAGAGCTGGCGGGGAGCTCGGCC 1154  
QY 809 AGGCAATGAGGCAITGTGAACGGCATGTGAGAGTGAACCCCAAGACCGACA 868  
Db 1155 CCGGCGCTCGGTGAGCTTCAACGGCATGAGTGAACCGCGCGGACCGCTCGGGGTGCA 1214  
QY 869 AGTTCCTGTCTGCGCCCTAGACAGAACAGCGTCTACGCGGCGGAGCGCCCGCCAAAG 928  
Db 1215 AGTCAACAGCGGCTCTCTGGGAGCGGCGGCGGAGCTCCCGGCGACACCGCCCGCCACGCGC 1274  
QY 929 AGGCTCTGACGCGGAGCTGGGCTGCTGTGGACCG--CACCAGCGGCTGTTCGCT 985  
Db 1275 CCGGACCGCCCGGAGCTGAGCTTCAACAGAACCTCGGTGAAGCTCTCTGGAAGCGG 1334  
QY 986 TCATCGCGCCCTGGAGGACAGAGGCTGTGACATCATCTTGGCGCGCTGCCAAGA 1045  
Db 1335 CCACCGACAGAACAGGCGCTCAAGAACTACGACGTCTTGGCGGAGCGGCGCAAGTCCCA 1394  
QY 1046 TCTGCGCACCGCCAGGTCAGATCGCATCTCTGGGTACCGGACAGGCGGCTACGAGA 1105  
Db 1395 CCGTACCGCCACCACTTACCGACAGAACTCGCCCGCGGACCGACTACTCTACT 1454  
QY 1106 AGTGTGAACGCGCATCGGACCAAGTACAAGGCGCGGCAAGGCGGTGTCAAGTCT 1165  
Db 1455 CGTCCAGCGCGGACCGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1514  
QY 1166 CGGCGCCCTGGGCGACATCTCACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1225  
Db 1515 CCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1568  
QY 1226 TGGAGCGCTGCGGCTGATCCAGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1285  
Db 1569 TCGGTACTTACCGAGTGGGCGTCTACGGCGCACTTACCAGTCAAAACCTGTGA 1628  
QY 1286 CTTCCACCGCGGCGGCTGTGACACCGTCAAGGAGGCGGTACCGGCTTCCACATGGCG 1345  
Db 1629 CTTCCGCGCTCGCGGAGAGATCACCACTCACTACTCTTTCGGCAACGTCACGCGGCG 1688  
QY 1346 CCTGAACCGCGGAGCTGGAGGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1405  
Db 1689 GC---AAGTGCACCATCGGTGACAGCTTCGCGGCTTACGACAGGCGTACCGCGCGG 1745  
QY 1405 GTGCCAGCGAGGTGTTGGGCGGCGGCTACCGGAGTGTGGCCAACTGATCAGCG 1465  
Db 1745 AGTCCGCTGACCGGCTCGCGGACCTTGGGACGAGCGGCTGCGCGGCACTTCAACGAG 1805  
QY 1466 AGGACTGTCTGTCGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1525  
Db 1806 TCCGCAAGTCAAGCGGAGTACCGGCACATCAAGTCTCTGTCGTCGTCGTCGTCGTCGTC 1865  
QY 1526 AGGCAAGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1585  
Db 1866 CTTGTCGCGGCGGCTTACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1925  
QY 1586 AGATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1645  
Db 1926 AGGACTGTGAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1985  
QY 1646 CTTCCGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1705  
Db 1986 AGTACCGGAGCGCTTGGCGGCTCAGCTGGACAGCTCCGCTCGGCGGCGGCTGGAAGA 2045  
QY 1706 GCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1765

Db 2046 TGGTCCAGCGATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2105  
QY 1766 CCACCTAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1825  
Db 2106 AGCCAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2165  
QY 1826 CCGTGTAGGCGCTCCACCACTCGACCTCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1885  
Db 2166 ACTGTCAACAGTGTGATGACCTACGACTTCTTCGGCGGCGGCGGCGGCGGCGGCGGCGG 2225  
QY 1886 AGGTGCTCGGCGGCTCCAAAGACCTCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1929  
Db 2226 CGCCCCACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2269

## RESULT 7

US-09-966-881-44  
; Sequence 44, Application US/09966881  
; Patent No. US20020120960A1  
; GENERAL INFORMATION:  
; APPLICANT: Seymour, Graham  
; Bird, Colin  
; Medina-Suarez, Rosybel  
; TITLE OF INVENTION: Genetic control of Fruit Ripening  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Zeneca Ag Products Inc.  
; STREET: 1800 Concord Pike  
; CITY: Wilmington  
; STATE: DE  
; COUNTRY: USA  
; ZIP: 19850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/966,881  
; FILING DATE: 28-Sep-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/242,860  
; FILING DATE: 29-Mar-1999  
; APPLICATION NUMBER: GB 9618862.8  
; FILING DATE: 10-SEP-1996  
; APPLICATION NUMBER: GB 9708366.1  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: PCT/GB97/02424  
; FILING DATE: 08-SEP-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hohenschutz, Liza D.  
; REGISTRATION NUMBER: 33,712  
; REFERENCE/DOCKET NUMBER: SEE 50183/UST  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (302) 886-1699  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 790 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: U-D111  
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:  
US-09-966-881-44

Query Match 4.8%; Score 94.2; DB 10; Length 790;  
Best Local Similarity 57.2%; Pred. No. 1.7e-12;  
Matches 210; Conservative 0; Mismatches 144; Indels 13; Gaps 2;  
QY 1 GCGCTGGACATCGTGATGTTGCTGCTGAGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCTG 60

Db 310 GGGATGACATAGTCTTTGTCGGCGCTGANATGCTCCCTGGAGTAAACCGGAGACTC 369  
QY 61 GCGATGAGTGTGGCTGCTCCCTATTGAGTGTCAAGCGCGCCCGCTGATGACC 120  
Db 370 GGTGATGTTCTTGGAGGACTGCCACCGGCCATGGCTGCAATGGACACAGAGTATGACT 429  
QY 121 ATTGCGCTCCCTACGACCACTAGCTGACCGCTGGGACACTCGGTGGTCTGTGACATC 180  
Db 430 ATAGTCCACGCTATGATCAGTACAAAGATGGTGGGATACAAATGCTCTGGTGAAGTTA 489  
QY 181 ATGGCGGAGAAGG-----TCGCTACTTCCACTCCATCAAGAAGGGGTGAC 228  
Db 490 AAGTTGGAATGAAATGAAAGANTCCGCTTCCACTCTATAAAGAAAGANTGAC 549  
QY 229 CGGTGTGATGATGACACCCCTGTTCTGCGCAAGTGTCTGGGCAAGCCGCTCCAG 288  
Db 550 AGGGTTTTCATGATGATCATCTTGTCTTGANAAAGTGTGGGAAAACTGGAAGAATG 609  
QY 289 CTGTACGCGCCCG-CTCGCGGCTGACTACTGTGACAAACACAAAGCGCTTCGCCCTGTT 347  
Db 610 ATATAGTCCCTGTCNCAGGAACNGATTAACAAACACAGCTAAGATTAACTTTTG 669  
QY 348 CTGCAAG 354  
Db 670 TTGCCAG 676

## RESULT 8

US-10-145-415-98  
; Sequence 98, Application US/10145415  
; Publication No. US20030068788A1  
; GENERAL INFORMATION:  
; APPLICANT: Buckel, Thomas Gunter  
; APPLICANT: Hammer, Philip Eugene  
; APPLICANT: Hill, Dwight Steven  
; APPLICANT: Ligon, James Madison  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Pachlatko, Johannes Paul  
; APPLICANT: Zirkle, Ross Eric  
; TITLE OF INVENTION: Methods and Compositions for Making Enamectin  
; FILE REFERENCE: SYN-117 109846.312  
; CURRENT APPLICATION NUMBER: US/10/145,415  
; CURRENT FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 60/291,149  
; PRIOR FILING DATE: 2001-05-16  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 98  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Streptomyces

## US-10-145-415-98

Query Match 4.4%; Score 86.6; DB 9; Length 1266;  
Best Local Similarity 44.2%; Pred. No. 8.6e-11;  
Matches 405; Conservative 0; Mismatches 509; Indels 3; Gaps 1;  
QY 1033 GCGCTCCCAAGATCTGGCCACCCCAAGTGCAGATCGCCATCTCTGGTACCGCA-- 1090  
Db 241 GCGCAGCCGCGCTCGCCCTCGACCCCGAGGAGGAGGAGCGCTCGCCCTCGCGCAGCGCACC 300  
QY 1091 -AGGCGCGCTACGAGAAGTGTGAACGGCATCGGCACCAAGTACAAAGGCGCGCGCAAG 1149  
Db 301 CTGATCGCTACGACAAGTGTGCTGGCCACCGCGCCGCAACCGCGCGCTTGGACATC 360  
QY 1150 GCGGTGTCAAGTTCTGGCGGCCCTGTGGCGCACATGTCACCGCGCGCGCGCTGACTCATG 1209  
Db 361 CCGGCGACCGCGCTGGCGCGCTGACACCACTCGCGCGCTCGCCCGCGCGCGCGCGCTG 420  
QY 1210 CTGTGCGCTTCGAGCGCTCGAGCGCTGATCGAGCTGATCGACCGCATGCACTACGGT 1269  
Db 421 CGCGCGCTCTGGCTCCCTCGCGCGCGGACACAGGCGCATCTGTGTATCGCGCGCGCGC 480

QY 1270 ACCGTGCCCCGTGGTAGCCTCCACCGCGCGCTGTGGTGCACACCGTCAAGGAGGCGCTCACC 1329  
Db 481 TGGATCGGCTTGGAGTTCGGCGCGCTCTTACGCGCGCGAGTGAACCGTCTGTC 540  
QY 1330 GGTTCACATGAGGCGCGCTGAACCGCGACAGCTGGAGGCTGACGCGGCTGTGTCACCG 1389  
Db 541 GAGCGCGCGCGCGCTGACCGCGCTGACCGCGCTGCTGGGCGCGCGAAGTCTGGCGGTGTGTCACC 600  
QY 1390 GCGCGCACCGTTCGCGCGTCCACGAGGTGTTTGGGCGCGCGCTACCCCGAGATGTTG 1449  
Db 601 GATCTGACCGCGACACGGCTCGCGCTTCCACTTCGCGCGCGCTTACCGGAGATGTC 660  
QY 1450 GCGAATCTCATCAGCCAGGACCTGCTCTGGTCCAAAGCGCGCGCGCGCGCGCGCGCTG 1509  
Db 661 GAGAGGCGCGCATGCTGCTCGCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCAC 720  
QY 1510 CTGGAGGAGTGTGTACGGCAAGGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1569  
Db 721 GATGTGCTGCGCGCATGCTG 780  
QY 1570 GTGCGCGTTCGCGGAGAAGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1629  
Db 781 GATCTGCGCGACCCCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCC 840  
QY 1630 CTGAAGCGCGTGTCCCGCTTCCGTGGAGGCAACGGCGCGCGCGCGCGCGCGCGCGCGCG 1689  
Db 841 GACCCGTACATCTACGCGCGCGGTGACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
QY 1690 ACCGCGCGCGCATGGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1749  
Db 901 CGGTGCGCGTGAACACTGGCG 960  
QY 1750 GCGACCGCGAAGGTGACCACTACAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1809  
Db 961 ATGCTGCGCGAGGACATCAGTACGACCGCATCCCGTACTTCTTCGCGCGCGCGCGCGCG 1020  
QY 1810 GTGGCGCTCAAGTGGCGCGTGGAGGCTTCACACCTCGACCTCGGAGAGACGCGCGTGGC 1869  
Db 1021 GTGGCATGAGTACTTCCGCGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080  
QY 1870 TCCACGCGAAGCGCAACGGTGTCTGCGCTCAAGACCTCGCGCTGCGCGCGCGCGCGCG 1929  
Db 1081 GCGGACGTGCGCGAGCGGAGTTCATCGCTTCTGCTGGCGCGCGCGCGCGCGCGCGCTC 1140  
QY 1930 TCG 1946  
Db 1141 GCGGCGCATGAACGTCAA 1157

## RESULT 9

US-09-893-519A-87  
; Sequence 87, Application US/09893519A  
; Publication No. US20030027243A1  
; GENERAL INFORMATION:  
; APPLICANT: ANADYS PHARMACEUTICALS, INC.  
; APPLICANT: THOMPSON, Craig  
; APPLICANT: MOORE, Jeffrey  
; APPLICANT: BUURMAN, Ed T.  
; APPLICANT: BRADLEY, John  
; APPLICANT: DESILVA, Tamara  
; APPLICANT: HARRIS, Sandra  
; APPLICANT: KOMARNITSKY, Svetlana  
; APPLICANT: MENDILLO, Marc  
; APPLICANT: MCCOY, Melissa  
; APPLICANT: SANDERSON, Karen  
; APPLICANT: HAQ, Tariq  
; APPLICANT: ZHU, Shuhao  
; APPLICANT: LONG, Fan  
; APPLICANT: DAVIDOV, Eugene  
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE  
; FILE REFERENCE: 0342/1G548-US2

[illegible]





Db	11361	ACCTCCCGCGGCTCTCGTGGCGGACCAACGACGCGCACCGGCTCTCAACAACCAACCACTAGC	11302
QY	1097	CCTACGAGAAGCTGGTCAACGCCCATCGGACCAAGTACAAGAGCGCGCGCAAGGCGGTGG	1156
Db	11301	TGATCGTCTGAGATCGACGAGGCCAACACCGGCATCCACCGCACTGTCATGGAGGTCC	11242
QY	1157	TCAAGTCTTCGGCGCCCTTGGCGGACATGCTACGCGCGCGCGGCGGCTCATGCTGCTGGTGC	1216
Db	11241	TGAAGCGGAAGCGGTGCACACCGCGGCTACTTCTCGCGGGCTGCCACGAGCTGGAGC	11182
QY	1217	CCTCGGCTTCGAGCGCCTCGGGCTCATCCAGCTGACGCCATGACATACGTTACCGTACCGTGC	1276
Db	11181	CGTACCGCGGCGACGCGACGCCCGCTCCGCGACACCAAGCGCTCGCGCGCGCGCTGC	11122
QY	1277	CCGTGGTAGCTCCACCGCGCGCTGGTGCACACCGTCAAGAGAGGCGCTCACCGGCTTCC	1336
Db	11121	TGTCCTTGCGACCGGACCGGCATCGGGACGACGACATCGCGGGTGCAGCACTGC	11062
QY	1337	ACATGGCGCCCTTGAAACCCGACAGCTGGACGAGGTGACGCCAGCGCTTGGCGGCA	1396
Db	11061	TGGCTCTCTGCGGACCCCGCGCGGAACTGACCGCGCGCACCGGACGCGGCGCCG	11002
QY	1397	CCGTGCGCGCTGCCAGCGAGTGTGTTGGGGCGCGCTACCCGAGATGTTGGCCAACT	1456
Db	11001	CCCCGCTCGGCGCCCC-----CAGACATCCAGCCCACTTGGAGCTCCCGA	10952
QY	1457	GCATCAGCAGGACCTGTCTGTGCCAGCGCGCCAGAGTGGAGGGCTGTCTGGAGG	1516
Db	10951	TGACCGCCCGCGCCCTTTCGCGCACCGCCCGCGGCGAACTGCGCGCACCCCGGAGCG	10892
QY	1517	AGTGGTGTACGCCAAGGCGCGTGGCCACCGCGCCCAAGAGGAGAGATCAAGTGTGCCG	1576
Db	10891	ATCTGGGGCGGGTCCACGCGCTGCGCGAGACCTCGCCGCGGGCGGCTCTGTGCGC	10832
QY	1577	TTGCGGAGAAGATCCCGCGCACTGCGCGCGTGTCTACGCGCCCAACACCTGAAGC	1636
Db	10831	CGCACGAGGCGGAAACGACCGCGCCCACTGCTCCGGCTCGCGCTACGGCAACA	10772
QY	1637	CGGTCTCGGCTCCGTGAGGGGCAACGGCGCGCGCGCCCAAGTGGGCAACCGCC	1696
Db	10771	GCCTTTCACCCCGCTGGAGAGGCGCCCGCACACTGGCGTGCAGCGGACGCTTCC	10712
QY	1697	CGGCATGGCGCGTGGGCGCGACACCCCTCGGGCCCTCGCGCGCGCGCGCCACCC	1756
Db	10711	GGCGCTCTCGCCCTGTTCGSGCAGTTCGCGGAGCTCCGACCGCGGTGAGACCGGCC	10652
QY	1757	CCAAGTGAACCTTACAAGCCGCGCTGCGCGCCACCGCAAGCCCAAGACCGCTGGCC	1816
Db	10651	CGCGCGGGCGGTACTTGAAGAACAACCTGTCTCCCGCTCGAACAGCGCGGCTCTTGACG	10592
QY	1817	TCAAGTGGCGGTGAGSCCTCCACCCTCGACCTCGGAGACGGCGTGGCTTCCAAAG	1876
Db	10591	CGGCGCTCGCCAGGAGGCCCTCTTCGCGTACAGCGTCGGCTCTACCCCGGCGCACCT	10532
QY	1877	GCA 1879	
Db	10531	GCA 10529	

RESULT 13

```

US-09-860-846-3/c
: Sequence 3, Application US/09860846
: Patent No. US2002016472A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding m
: FILE REFERENCE: 600 438US1
: CURRENT APPLICATION NUMBER: US/09/
: PRIOR FILING DATE: 2001-05-18
: CURRENT APPLICATION NUMBER: 09/105,5

```

QY	1637	CCGTCGCGCTCCGCTGGAGGCAACGGCGCGCGCCGCCAAGGTGCGCACCAACCGGCC	1696	Db	11988	TGATGTCGAGATCAGAGGCCACACCGGATCATCCGCGACCTCGTCATGGAGTCC	11929
Db	11458	GCCCCCTTACCCCGCTGAGGAGGCCCGCCAGACCTGGCGTGACCGGAGCGCTTCC	11399	QY	1157	TCAAGTTCCTGGCGCCCTGGCGGCACATGCTCACCGCGCGCGCCGACCTTCATGCTGGTGC	1216
QY	1697	CCGCGATGGCGGTGGCGCGGACACCCCTCGGGGCCCTCGCCCGCGCGCCACCC	1756	Db	11928	TGAAGCGCAAGCGGTGCACACCGCGGCTACTTCTCGCGGGCTGCCACGAGCTGGAGC	11869
Db	11398	GGCGCTCTCGCCCTGTTCGGGAGTCCCGAGCTCCGACCGCGGTTCGAGACCGGCC	11339	QY	1217	CCTCGGCTTCGAGCCCTGGCGGCTGATCCAGCTGACGCCATGACTACGCTACCGTGC	1276
QY	1757	CCAGGTGACCACTACAGCCCGCCCTGGCGCCCGCCACCGCAAGCCCAAGACCGCTGGCC	1816	Db	11868	CGTACCGGGCGGACCGGACCGCGGCTGCCGACACCGCAAGCGCTTCGCGCGCGGTGC	11809
Db	11338	CGCGCGGGCGTACTGGAAGAACACCTGCTCCCGCTCAACAGCGCGGCTTCGAGG	11279	QY	1277	CGGTGAGTCTCCACCGCGGCTGTCGACACGCTCAAGGAGGGCGTCAACGGCTTCC	1336
QY	1817	TCAAGTGGCGGTGAGGCGCTCCACCTCGACCTCGAGAGCGCGCTGCGCGCAAG	1876	Db	11808	TGTCCTTGCAGCCGACACCGCTCGGCGAGACACATCCCGGGTTCGCGACCTGC	11749
Db	11278	CGGCGCTCGCAGGAGCGCGCTTCCCGTACAGCGTCCGCTTACCCCGCGCGGCT	11219	QY	1337	ACATGGCGCCCTGAACCCCGACAGCTGGAGGCTGACGCGCAGCGCTTGCGCGCA	1396
QY	1877	GCA 1879		Db	11748	TGGTCTCTGGCGGACCGCGCGGCGGAACTGACCGCGCGCACCGCGACAGCGCGCG	11689
Db	11218	GCA 11216		QY	1397	CCGTGCGCGCTGCCAGCGAGTGTTCGCGGGCGCGCTACCCCGAGATGGTGGCCACT	1456
RESULT 14							
US-09-836-821-3/c							
; Sequence 3, Application US/09836821							
; Publication No. US20030087405A1							
; GENERAL INFORMATION:							
; APPLICANT: Sherman, D.H.							
; APPLICANT: Liu, H.							
; APPLICANT: Xue, Y.							
; APPLICANT: Zhao, L.							
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin							
; FILE REFERENCE: 600.438US1							
; CURRENT APPLICATION NUMBER: US/09/836,821							
; CURRENT FILING DATE: 2001-04-17							
; PRIOR APPLICATION NUMBER: 09/105,537							
; PRIOR FILING DATE: 1998-06-26							
; NUMBER OF SEQ ID NOS: 43							
; SOFTWARE: FastSeq for Windows Version 3.0							
; SEQ ID NO 3							
; LENGTH: 13613							
; TYPE: DNA							
; ORGANISM: Streptomyces venezuelae							
US-09-836-821-3							
Query Match 4.1%; Score 81; DB 9; Length 13613;							
Best Local Similarity 42.7%; Pred. No. 1.5e-09;							
Matches 488; Conservative 0; Mismatches 645; Indels 10; Gaps 1;							
QY	737	CGCCCAACTACGCGACCGAGATCGTCCGATGCGCGCGCGCGGTGGAGCTGGACACCG	796	Db	12348	CGGTGACG	12289
QY	797	TCATCCGCGCGCAAGGCGATTGAGGCGATGTGAACGCGCATGGACATGAGGATGGAAC	856	Db	12288	CCAAGCGGTCAACGCGCTTGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	12229
QY	857	CCAAGCGCGCAAGTTCCTGTCGCGCCCTACGACCAAGACAGCGTCAAGCGCGCAAGG	916	Db	12228	CCGCGATCGCGCGCTCCCAACTTCGGCTTCGACCTGCGCGCGCGCGCGCGCGCGCG	12169
QY	917	CCGCGCGCGCAAGGCGCCCTCGAGCGCGAGCTGGCGCTTCGCTGTGGACCCCGCCCG	976	Db	12168	GGACCAACGCGAAGATGAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	12109
QY	977	TGTCGCGCTTCATCGCGCGCTCGAGGAGGAGGCTGGACATCATCTTGGCGCGCC	1036	Db	12108	TTCCGAGGTATCGACCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	12049
QY	1037	TGCCCAAGATCTCGCGCACCCCGCAAGGTGAGATCGCATCTCTCGGTACCGGCAAGCG	1096	Db	12048	ACCTCCCGCGGTCTCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	11989
QY	1097	CCTACGAGAGCTGTTGTAACGCCATTCGCGACCAAGTACAGAGGCGCGCGCGCGCG	1156				

RESULT 15  
US-09-861-289-3/c  
; Sequence 3, Application US/09861289  
; Patent No. US20020110897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/861,289  
; CURRENT FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: 09/105,537  
; PRIOR FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:49:45 ; Search time 34.5277 Seconds  
(without alignments)  
2512.365 Million cell updates/sec

Title: US-09-980-771a-5  
Perfect score: 3370  
Sequence: 1 ALDIYVAAEAPWPSKTGL.....SASKTSAAKPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1:	/SID22/gcgdata/genseq/genseq-emb1/AA1980.DAT:*
2:	/SID22/gcgdata/genseq/genseq-emb1/AA1981.DAT:*
3:	/SID22/gcgdata/genseq/genseq-emb1/AA1982.DAT:*
4:	/SID22/gcgdata/genseq/genseq-emb1/AA1983.DAT:*
5:	/SID22/gcgdata/genseq/genseq-emb1/AA1984.DAT:*
6:	/SID22/gcgdata/genseq/genseq-emb1/AA1985.DAT:*
7:	/SID22/gcgdata/genseq/genseq-emb1/AA1986.DAT:*
8:	/SID22/gcgdata/genseq/genseq-emb1/AA1987.DAT:*
9:	/SID22/gcgdata/genseq/genseq-emb1/AA1988.DAT:*
10:	/SID22/gcgdata/genseq/genseq-emb1/AA1989.DAT:*
11:	/SID22/gcgdata/genseq/genseq-emb1/AA1990.DAT:*
12:	/SID22/gcgdata/genseq/genseq-emb1/AA1991.DAT:*
13:	/SID22/gcgdata/genseq/genseq-emb1/AA1992.DAT:*
14:	/SID22/gcgdata/genseq/genseq-emb1/AA1993.DAT:*
15:	/SID22/gcgdata/genseq/genseq-emb1/AA1994.DAT:*
16:	/SID22/gcgdata/genseq/genseq-emb1/AA1995.DAT:*
17:	/SID22/gcgdata/genseq/genseq-emb1/AA1996.DAT:*
18:	/SID22/gcgdata/genseq/genseq-emb1/AA1997.DAT:*
19:	/SID22/gcgdata/genseq/genseq-emb1/AA1998.DAT:*
20:	/SID22/gcgdata/genseq/genseq-emb1/AA1999.DAT:*
21:	/SID22/gcgdata/genseq/genseq-emb1/AA2000.DAT:*
22:	/SID22/gcgdata/genseq/genseq-emb1/AA2001.DAT:*
23:	/SID22/gcgdata/genseq/genseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	42.7	606	13	Granule-bound star
2	1397.5	41.5	609	19	Oryza sativa starch
3	1390	41.2	527	21	Arabidopsis thalia
4	1390	41.2	563	21	Arabidopsis thalia
5	1390	41.2	610	21	Arabidopsis thalia
6	1390	41.2	610	23	Arabidopsis thalia
7	1353.5	40.2	533	19	Herbicideally activ
8	1344.5	39.9	637	19	Zea mays waxy gene
9	1331.5	39.5	502	13	Protein encoded by
10	1194	35.4	238	22	Central fragment f
					A granule bound st

11	904	26.8	534	21	AA49307	Wheat starch synth
12	898	26.6	792	23	ABB92160	Herbicideally activ
13	891.5	26.5	647	20	AA090004	Wheat starch solub
14	882.5	26.2	649	19	AAW38218	Maize starch synth
15	881	26.1	671	19	AAW23937	Wheat soluble star
16	881	26.1	756	21	AA50818	Wheat soluble star
17	873.5	25.9	626	15	AA51231	Soluble rice starch
18	870.5	25.8	625	23	ABB93595	Herbicideally activ
19	861.5	25.6	539	19	AAW56491	Zea mays starch sy
20	861.5	25.6	583	19	AAW56488	Herbicideally activ
21	852.5	25.3	583	19	AAW70894	Zea mays soluble s
22	847	25.1	802	23	AAU99845	Maize starch solub
23	847	25.1	813	23	AAU99844	Modified barley st
24	846	25.1	812	23	AAU99847	Barley cultivar Mo
25	845	25.1	804	19	AAW70892	Barley line MK827
26	844	25.0	798	21	AAW70892	Maize starch solub
27	842	25.0	799	21	AAW70892	Wheat starch synth
28	842	25.0	799	21	AAW70892	Wheat starch synth
29	841	25.0	799	21	AAW70892	Wheat starch synth
30	837	24.8	698	19	AAW56487	Wheat starch synth
31	825.5	24.5	812	23	AAU99846	Wheat granule-boun
32	825	24.5	597	21	AAW37567	Zea mays soluble s
33	767	22.8	466	23	AAW51865	Barley line 292 st
34	726	21.5	459	17	AAW51865	Wheat starch synth
35	701.5	20.8	669	19	AAW70893	Soluble starch syn
36	701.5	20.8	669	19	AAW70893	Rice starch synth
37	669	19.9	483	23	AB54012	Soluble starch solub
38	647.5	19.2	476	23	AB54012	Zea mays soluble s
39	580	17.2	477	13	AAW25462	Lactococcus lactis
40	580	17.2	477	19	AAW53890	Streptococcus poly
41	568.5	16.9	677	17	AAW53890	glgA. Escherichia
42	568.5	16.9	1230	18	AAW17785	E. coli glycogen s
43	568.5	16.9	1230	21	AAW17785	Soluble starch syn
44	561.5	16.7	1025	23	AB54012	Potato tuber solub
45	546	16.2	495	19	AAW70885	Potato starch synt
						Herbicideally activ
						Protein encoded by

## ALIGNMENTS

RESULT 1  
AAR25476  
ID AAR25476 standard; Protein; 606 AA.  
XX  
AC AAR25476;  
XX  
XX  
DT 15-JAN-1993 (first entry)  
XX  
DE Granule-bound starch synthase of potato.  
XX  
KW GBSS gene; amylopectin; amylose production;  
KW inhibition; promoter; antisense construct.  
XX  
OS Solanum tuberosum.  
XX  
PN WO9211376-A.  
XX  
PD 09-JUL-1992.  
XX  
PF 20-DEC-1991; 91WO-SE00892.  
XX  
PR 21-DEC-1990; 90SE-0004096.  
XX  
XX (AMYL-) AMYLOGENE HB.  
XX  
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;  
XX  
DR WPI; 1992-250096/30.  
XX  
DR N-PSDB; AAQ26404.  
XX  
PT Modifying potatoes to form amylopectin starch - using an  
PT anti-sense construct to inhibit granule-bound starch synthase

PS Claim 6; Page 28-33; 46pp; English.

XX A genomic library in ENBL3 was prepared using leaves of the potato  
CC Bintje. The library was screened with cDNA clones for the 5' and 3'  
CC ends of the GBSS gene. A full-length clone of potato GBSS gene was  
CC identified (wx311) and isolated from the genomic library. The gene  
CC contained 12 introns. The amino acid sequence was deduced from it.  
CC Three fragments of the full-length sequence (5'-end, middle and  
CC 3'-end sequences) were characterised as suitable for use in novel  
CC antisense constructs to suppress amylose formation in potatoes.  
CC See also Q26400-3.

XX Sequence 606 AA;

Query Match 42.7%; Score 1440; DB 13; Length 606;  
Best Local Similarity 53.5%; Pred. No. 9.2e-105;  
Matches 292; Conservative 67; Mismatches 145; Indels 42; Gaps 10;  
QY 2 LDIIVMAAEVAPWSKTGGIGDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60  
DB 81 MNLIPVTEGVPSKTGELGDLGGLPPALAAARGHRVMTISPRYDQYKDAWDTGVAVEVK 140  
QY 61 MG---EKVRYHSIKKGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117  
DB 141 VGDSTIEIVRFHCYKRGVDRVDFHPFLEKVMGTGSKYIYGPAGLGYLDNHRFSLLC 200  
QY 118 KAAIEAARVLPE-----GP-GEPCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLA 169  
DB 201 QAALEAPKVLNLSNYSFGPYGDEVLFANDWHFALIPCYLKSVMYQSRGIYLNKAVAF 260  
QY 170 IHNTAFQGRMWEAEFKTKLPQAFDKLAFSDGYAKVYTEATPMEDEKPPITGKTYYKI 229  
DB 261 IHNTAFQGRMWEAEFKTKLPQAFDKLAFSDGYAKVYTEATPMEDEKPPITGKTYYKI 229  
QY 230 NMLKGGIITADKLVTVPNVAETIADAAGVVELDTVIRAKGIEGIVNGMDIEENPKTD 289  
DB 303 NMKAGILESHRVTVSPYIAQELYSADVKGVELDSVLRKTCITGIVNGMDTQEWNPATD 362  
QY 290 KELSAPYQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFIRGLEQKGVDTIILAAALPKI 349  
DB 363 KYTDVKYDITVNDKAPLKEALQAAVGLPVKPKIPLIGFIRGLEQKGVDTIILAAALPKI 422  
QY 350 LATPKVQIAILGTGKAAAYEKLNAIGTKYKGRAGVYVTSAPLAHMLTAGADPMLVPSRF 409  
DB 423 IGL-DVQIVVLTGKKEFEQTELELVLPNKAQVAKENVPPLAHMLTAGADPMLVPSRF 481  
QY 410 EPCGLIQLHAMHYGTVPVASTGGIVDTYKGVTFGFMGALNP--DKLDEADADALAATV 467  
DB 482 EPCGLIQLHAMHYGTVPVASTGGIVDTYKGVTFGFMGALNP--DKLDEADADALAATV 467  
QY 468 RRASEVFAGGRYPENVANCISQDLSWSKPAQKWEGLL-----BEVYVKGGVATFAKK 519  
DB 542 ARALVYGTILAFAMINKMSELSKEPAKKWETLLLGIGASSEPGVEGE-EIAPLAK 600  
QY 520 EEIKVP 525  
DB 601 ENVATP 606

RESULT 2  
AAW56485  
ID AAW56485 standard; Protein; 609 AA.

XX AAW56485;

AC 11-SEP-1998 (first entry)

DT Oryza sativa starch (bacterial glycogen) synthase.

XX SER; starch-encapsulating region; fusion vector;

KW starch synthase; bacterial glycogen.

XX Oryza sativa.

OS

XX WO9814601-A1.  
PN 09-APR-1998.  
XX 30-SEP-1997; 97WO-US17555.  
PF 30-SEP-1996; 96US-0026855.  
PR (EXSE-) EXSEED GENETICS LLC.  
XX Guan H, Keeling P;  
PI WPI; 1998-240100/21.  
XX N-PSDB; AAV29753.  
DR Hybrid polypeptide comprising starch-encapsulating region and  
XX protein - useful for, e.g. producing protein(s) resistant to  
PT degradation by stomach acids  
PT Example 2; Page 32; 156pp; English.  
XX The sequence is that of starch (bacterial glycogen) which is  
CC encoded by the waxy gene. It can be used in the production of a hybrid  
CC polypeptide comprising a starch-encapsulating region (SER) fused  
CC to a payload protein. The hybrid polypeptide can be used to make  
CC modified starches comprising the payload protein, selected from,  
CC e.g. hormones, growth factors, antibodies, enzymes, dyes,  
CC immunoglobulins, etc. The modified starch can also be used  
CC to provide grain feeds enriched in amino acids. By encapsulating  
CC the payload protein in starch, it is more resistant to  
CC degradation by stomach acids.  
XX Sequence 609 AA;  
SQ Query Match 41.5%; Score 1397.5; DB 19; Length 609;  
Best Local Similarity 51.4%; Pred. No. 2.1e-101;  
Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;  
QY 2 LDIIVMAAEVAPWSKTGGIGDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 60  
DB 83 MNVVFVGAEMAPWSKTGGIGDVTGGLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDI- 142  
QY 61 ---MGEKVRVYFHSIKKGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117  
DB 143 VADRYERVRFFHCYKRGVDRVDFHPFLEKVMGTGSKYIYGPAGLGYLDNHRFSLLC 202  
QY 118 KAAIEAARVL-----PFGP---GEPCVFVANDHSAVLPVLLKDEYQPKGQFTKAKSVLA 169  
DB 203 QAALEAPKVLNLSNYSFGPYGDEVLFANDWHFALIPCYLKSVMYQSRGIYLNKAVAF 262  
QY 170 IHNTAFQGRMWEAEFKTKLPQAFDKLAFSDGYAKVYTEATPMEDEKPPITGKTYYKI 229  
DB 263 IHNTAFQGRMWEAEFKTKLPQAFDKLAFSDGYAKVYTEATPMEDEKPPITGKTYYKI 229  
QY 230 NMLKGGIITADKLVTVPNVAETIADAAGVVELDTVIRAKGIEGIVNGMDIEENPKTD 289  
DB 306 NMKAGILEADRVTVSPYIAQELYSADVKGVELDSVLRKTCITGIVNGMDTQEWNPATD 365  
QY 290 KELSAPYQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFIRGLEQKGVDTIILAAALPKI 349  
DB 366 KYTDVKYDITVNDKAPLKEALQAAVGLPVKPKIPLIGFIRGLEQKGVDTIILAAALPKI 425  
QY 350 LATPKVQIAILGTGKAAAYEKLNAIGTKYKGRAGVYVTSAPLAHMLTAGADPMLVPSRF 409  
DB 426 M-QEDVQIVVLTGKKEFEQTELELVLPNKAQVAKENVPPLAHMLTAGADPMLVPSRF 484  
QY 410 EPCGLIQLHAMHYGTVPVASTGGIVDTYKGVTFGFMGALNP--DKLDEADADALAATV 467  
DB 485 EPCGLIQLHAMHYGTVPVASTGGIVDTYKGVTFGFMGALNP--DKLDEADADALAATV 467  
QY 468 RRASEVFAGGRYPENVANCISQDLSWSKPAQKWEGLL-----VATAK 518  
XX

Db 545 KRAIKVVGTPAYEEMVRNMCNODLSWKGPAKNWENVL--LGLGVAGSAPGEGDEIAPLA 602  
QY 519 KEEIKVP 525  
Db 603 KENVAAP 609

RESULT 3  
AAG04668  
ID AAG04668 standard; Protein; 527 AA.  
AC AAG04668;  
XX  
XX  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 782.  
DE  
DE  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142390.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.





PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.





QY	287	KTDKFLSAPYDQNSVYAGKAAKEALQAEGLGVPDPTAPLFAFGRLEEQKGVDTILAAL	346
DB	364	STDYIDIKYDITVTDAKPLIKALQAAVGLPVDROVPVIGTIGRLEEQKGSLLVEAI	423
QY	347	PKILATPKVQIALLGTGKAAVEKIVNAIGTKYKGRAGKGVVKSAPLAHMLTAGADFMVLP	406
DB	424	SKFMGL-NVQMVLITGKGRKMEAIILEEFKPGKAVGAKFNVPPLAHMITAGADFIIVP	482
QY	407	SRFPCGLIQLHAMHYGTVPVVASTGGLVDIVKGVGTGFHMGALNP--DKLDEADADALA	464
DB	483	SRFPCGLIQLHAMHYGTVPVVASTGGLVDIVKGVGTGFHMGALNP--DKLDEADADALA	542
QY	465	ATVRRASEVFAGGRYPENVANCISQDLSWSKPAQKWEGLLEVVYKGGVATAKKEIKV	524
DB	543	KAVTRAVAVYGTSAQEMVKNMCDQDFSWKGPRLWEKVLVLSLVA-GSEAGTEGEET-A	600
QY	525	PVAEK 529	
DB	601	PLAKE 605	
RESULT 6			
ABE91229	standard; Protein; 610 AA.		
XX	AC	ABE91229	
XX	AC	ABE91229	
XX	AC	ABE91229	
DT	31-MAY-2002	(first entry)	
XX	DE	Herbicide active polypeptide SEQ ID NO 440.	
XX	XX	Herbicide; plant; agriculture; herbicide.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200210210-A2.	
XX	PD	07-FEB-2002.	
XX	XX	28-AUG-2001; 2001WO-EP09892.	
XX	XX	28-AUG-2001; 2001WO-EP09892.	
XX	PA	(PARB ) BAYER AG.	
XX	PI	Tietjen K, Weidner M;	
XX	XX	WPI; 2002-269010/31.	
XX	XX	Identifying plant target proteins for herbicidally active compounds,	
PT	PT	comprising aligning and comparing nucleic acid or amino acid sequences	
PT	PT	from plant with nucleic acid or amino acid sequences from non-plant	
PT	PT	organisms -	
XX	XX	Claim 5; SEQ ID NO 440; 261pp + Sequence Listing; English.	
XX	XX	The invention relates to identifying target proteins	
CC	CC	(ABB90790-ABB94016) for herbicidally active compounds, comprising	
CC	CC	aligning and comparing nucleic acid or amino acid sequences from plant	
CC	CC	with nucleic acid or amino acid sequences from non-plant organisms using	
CC	CC	with nucleic acid or amino acid sequences from non-plant organisms using	
CC	CC	suitable search parameters, where plant sequences having an E-value	
CC	CC	greater by a factor of 3 than the E-value of most similar non-plant	
CC	CC	sequences are selected. The polypeptides or nucleic acids encoding them	
CC	CC	are useful for identifying modulators. The identified modulators are	
CC	CC	useful as herbicides.	
XX	XX	Sequence 610 AA;	
SQ	XX	Query Match 41.2%; Score 1390; DB 23; Length 610;	
		Best Local Similarity 52.7%; Pred. No. 8.2e-101;	
		Matches 287; Conservative 72; Mismatches 146; Indels 40; Gaps 12;	
QY	2	LDIVWAAAEVAPWSKTGGLGVDIVKGVGTGFHMGALNP--DKLDEADADALA	600

DB	84	MSVIFGAEVGSWSTGGLGVDIVKGVGTGFHMGALNP--DKLDEADADALA	143
QY	61	MGEK---VRYFHSIKKGVHVRWIDHPWELAKYWGKSGKLYGPRSGADYLDNHHKRALFC	117
DB	144	VGDKYENVRFFHCYKRGVDREVDHPFLAKVVGKSGKLYGPRSGADYLDNHHKRALFC	203
QY	118	KAAIEAARVLPF-----GP-GEDCVFVANDHWSALVPVLKDEYQKQGTAKKSVLA	169
DB	204	QAALAPQVNLNLSKYPGSGPYGVDVFWFVANDHWSALVPVLKDEYQKQGTAKKSVLA	263
QY	170	INIAFOGRMEEAFKDKLP--QAADFCLAFSDGAKVYVTEATPMEDEKPPILTKTY	226
DB	264	INIAFOGRMEEAFKDKLP--QAADFCLAFSDGAKVYVTEATPMEDEKPPILTKTY	303
QY	227	KYNWLGKGIITAAADLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEENP	286
DB	304	RKINWMAALEAHRLVTVSPNYATEIAADAAGVVELDTVIRAKGIEGVNGMDIEENP	363
QY	287	KTDKFLSAPYDQNSVYAGKAAKEALQAEGLGVPDPTAPLFAFGRLEEQKGVDTILAAL	346
DB	364	STDKYIDIKYDITVTDAKPLIKALQAAVGLPVDROVPVIGTIGRLEEQKGSLLVEAI	423
QY	347	PKILATPKVQIALLGTGKAAVEKIVNAIGTKYKGRAGKGVVKSAPLAHMLTAGADFMVLP	406
DB	424	SKFMGL-NVQMVLITGKGRKMEAIILEEFKPGKAVGAKFNVPPLAHMITAGADFIIVP	482
QY	407	SRFPCGLIQLHAMHYGTVPVVASTGGLVDIVKGVGTGFHMGALNP--DKLDEADADALA	464
DB	483	SRFPCGLIQLHAMHYGTVPVVASTGGLVDIVKGVGTGFHMGALNP--DKLDEADADALA	542
QY	465	ATVRRASEVFAGGRYPENVANCISQDLSWSKPAQKWEGLLEVVYKGGVATAKKEIKV	524
DB	543	KAVTRAVAVYGTSAQEMVKNMCDQDFSWKGPRLWEKVLVLSLVA-GSEAGTEGEET-A	600
QY	525	PVAEK 529	
DB	601	PLAKE 605	
RESULT 7			
AAW56484	standard; Protein; 533 AA.		
XX	AC	AAW56484	
XX	AC	AAW56484	
DT	11-SEP-1998	(first entry)	
XX	XX	Zea mays waxy gene glucosyl transferase.	
DE	XX	SER; starch-encapsulating region; fusion vector.	
XX	XX	Zea mays.	
XX	XX	WO9814601-A1.	
XX	XX	09-APR-1998.	
XX	XX	30-SEP-1997; 97WO-US17555.	
XX	XX	30-SEP-1996; 96US-0026855.	
XX	XX	(EXSE-) EXSEED GENETICS LLC.	
XX	XX	Guan H, Keeling P;	
XX	XX	WPI; 1998-240100/21.	
XX	XX	N-PSDB; AAV29752.	
XX	XX	Hybrid polypeptide comprising starch-encapsulating region and	
PT	PT	protein - useful for, e.g. producing protein(s) resistant to	
PT	PT	degradation by stomach acids	
XX	XX		



Db	480	FEPCGLIQLQGMRYCTPCACASTGGLVDIIIEGKTGFHMGRLSDVNCVVEPADVKKVATT	539
QY	467	VRRASEVAGGRYPEMVANCISQDLSWSKPAQKWEGLLEE--VYVGKGV-----ATAKK	519
Db	540	LQRAIKVGTAYEEMVRNCMIQDLSWKGPAKNWNLVLSLVGAGGEPGVEGETAPLAK	599
QY	520	EEIKVP	525
Db	600	ENVAAP	605
RESULT 9			
ID	AAR25474	standard; Protein; 502 AA.	
XX	AAR25474;		
DT	15-JAN-1993	(first entry)	
DE	Central fragment from potato GBSS.		
XX	Granule-bound starch synthase; amylopectin; amylose production; inhibition.		
OS	Solanum tuberosum.		
PN	W09211376-A.		
XX	09-JUL-1992.		
PF	20-DEC-1991;	91WO-SE00892.	
XX	21-DEC-1990;	90SE-0004096.	
XX	(AMY1-) AMYLOGENE HB.		
PI	Hofvander P, Persson PT, Tallberg A, Wikstrom O;		
DR	WPI; 1992-250096/30.		
DR	N-PSDB; AAQ26401.		
XX	Modifying potatoes to form amylopectin starch - using an anti-sense construct to inhibit granule-bound starch synthase		
PS	Claim 1; Page 21; 46pp; English.		
CC	A genomic library in EMBL3 was prepared using leaves of the potato Bintje. The library was screened with cDNA clones for the 5' and 3' ends of the GBSS gene. A full-length clone of potato GBSS gene was identified (wx311). A BglII-SpeI fragment ("m") of the clone was found to contain the central region of the gene and was cloned in pUC13 to give pSm. Restriction of pSm with NsiI and HpaII gives fragment II which was cloned in pJRD184 to give pJRDmitt. Further restriction of pJRDmitt with HpaII-SstI gives a 2549bp fragment comprising exons and introns from the middle of the gene. The fragment can be used to make antisense constructs to suppress amylose formation in potatoes. See also Q26400-4.		
SQ	Sequence 502 AA;		
Query Match 39.5%; Score 1331.5; DB 13; Length 502;			
Best Local Similarity 55.6%; Pred. No. 2.5e-96;			
Matches 269; Conservative 55; Mismatches 125; Indels 35; Gaps 9;			
QY	2	LDIVMVAEYAPNSKGTGGLDVTGGLPIELVKGHRVMTIAPRYDQYADAWDTSVVVDI-	60
Db	39	MNLIFVTEGVPWSKGTGGLDVLGGLPALAAARGHRVMTISPRDQKDAWDGTGVAEVK	98
QY	61	MG---EKRVYFHSIKKGVHRWIDHPFLAKVWGKTSKLYGPRSGADYLDNKKRFALFC	117
Db	99	VGDSIEIVRFHCYKRGVDRVFDHPFMELEKVGKTSKIYGPAGLDYLDNLRFSILC	158
QY	118	KAATEAARVLFF-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKQFTKAKSVLA	169

CC with these grains. The recombinant nucleic acid sequence is used to  
 CC target polypeptides of interest to starch grains.  
 XX

SQ Sequence 238 AA;

Query Match 35.4%; Score 1194; DB 22; Length 238;  
 Best Local Similarity 99.2%; Pred. No. 5.9e-86;  
 Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 416 QHAMHYGVVPVASTGGLVDVTKGVTGFGHMGALNPKLDLADADALAATVRRASEVEFA 475  
 DB 3 RVHAMHYGVVPVASTGGLVDVTKGVTGFGHMGALNPKLDLADADALAATVRRASEVEFA 62  
 QY 476 GGRYPFWANCISQDLSSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKETPGDLP 535  
 DB 63 GGRYPFWANCISQDLSSKPAQKWEGLLEEVYKGGVATAKKEIKVPVAKETPGDLP 122  
 QY 536 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSPAAATPKVTYKPA 595  
 DB 123 AVSYAPNTLKPVSASVEGNGAAAPKVGTTAPAMGAWRATTPSGPSPAAATPKVTYKPA 182  
 QY 596 PATAKPKTAGLKLAGEASTTSTSENGASNGNGASASKTSAAPLYSAAATRKSA 651  
 DB 183 PATAKPKTAGLKLAGEASTTSTSENGASNGNGASASKTSAAPLYSAAATRKSA 238

RESULT 11

AAB49307  
 ID AAB49307 standard; Protein; 534 AA.

XX AAB49307;

AC AAB49307;

DT 01-MAR-2001 (first entry)

XX Wheat starch synthase GRSS protein.

DE Wheat starch synthase GRSS protein.  
 KW Wheat; starch synthase; SSII; SSIII; starch content; starch synthesis;  
 KW food product; adhesive.

XX Triticum aestivum.

XX WO200066745-A1.

XX 09-NOV-2000.

XX 28-APR-2000; 2000WO-AU00385.

XX 29-APR-1999; 99AU-0000052.

XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.

XX (GOOD-) GOODMAN FIELDER LTD.

XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX Morell M, Li Z, Rahman S, Appels R;

XX WPI; 2000-647602/62.

XX Nucleic acid molecules encoding wheat starch synthase (WST)-I and  
 XX WST-II, useful in modifying plant starch content and/or composition -  
 XX Example 15; Fig 9; 21pp; English.

XX The present invention relates to novel protein and coding sequences from  
 CC wheat. The proteins are wheat starch synthases, designated SSII and  
 CC SSIII. These can be used in the modification of plant starch content or  
 CC composition, and to screen plants to identify mutations which affect  
 CC starch content and composition. The starch can then be used in food  
 CC products, such as flour, and in films, coatings, adhesives, building  
 CC materials and packaging materials.

XX Sequence 534 AA;

SQ Query Match

26.8%; Score 904; DB 21; Length 534;

Best Local Similarity 32.5%; Pred. No. 1.4e-62;  
 Matches 213; Conservative 48; Mismatches 103; Indels 292; Gaps 9;

QY 7 VAAEVAPWSKTGGLGDTGTLPIELVKRGHRYMTIAPRDQYADADNTSVVDIMGEKVR 66  
 DB 2 VAAEMAPWSKTGGLGDLGLPPMAAANGHRVMVISPRYDQYKDAWDT----- 49  
 QY 67 YFHSIKKGVRHVMIDHPWFLAKWGTGSKLYGPRSGADYLDNHHKFAFLCKAAIAAARV 126  
 DB 50 -----LEKVRGKTKEKIVGPDAGTIDEDNQRFSLCQAALFVPI 90  
 QY 127 L-----PF----- 129  
 DB 91 LNLNNPFFCIHNISYQGRFSFDDPAQLNLPDRFKSSFDFIDGYDKPVEGRKISVVSEI 150  
 QY 130 -----GP-GEDCVFVANDHWSALVPVLLKDYQPK 158  
 DB 151 KVVDKYVRYRPHCYKRGVDRVFDHPGSGPYGEDVFCVNDWHTGLLACYLKSNYQSN 210  
 QY 159 GQTKAKSVLAIHNIAPQGRMWEAPKDTKLPAAPDKLAFSDGYAKVYVTEATPMEDEK 218  
 DB 211 GIYRAAKVA----- 219  
 QY 219 PPLTKYTKKINWLKGGIITADKLVTSYNYATEIAADAAGVVELDVIIRAKGIEGIVNG 278  
 DB 220 -----NWMKAGILQADKLVTSYNYATEIAADAAGVVELDVIIRAKGIEGIVNG 268  
 QY 279 MDIEEWNPKTDFLSAPYDONSAYAKAAK----- 309  
 DB 269 MDVSEWDPKTKFLAVNYDITTALEKGLNKLKEEGKALNKLKEEDVQIVLGTGKKKFE 328  
 QY 310 -----EALQAEGLPVDTPAPLFAFTRLEOKGVDTILA 344  
 DB 329 RLLKSTIEKFPKSVRAVVRFNAPLAEALQAEGLPVDKRVPLVAFTRLEOKGVDTILA 388  
 QY 345 ALPKILATPKVOIAILGTGKAAAYEKLNVNAIGTKYGRKGVVYKFSAPLAHMLTAGADFML 404  
 DB 389 SIPEI-----HOMMAGADVLA 404  
 QY 405 VPSREFPCGLIQHAMHYGVTPVPAVSTGGLVDVTKGVTGFGHMGALNPD--KLDEADADA 462  
 DB 405 VTSREFPCGLIQLOQMRVGTACACASTGGLVDVTKGVTGFGHMGALNPD--KLDEADADA 462  
 QY 463 LAATVRRASEVFPAGGRYPE-----MVANCISQDLSSKPAQKWEGLLEEV 507  
 DB 465 VVTLKRAVKVGVTPAYHEAPLAMEVNAAPVYKNCMIQDLSSKPAQKWEGLLEEV 520

RESULT 12

ABB92160

ID ABB92160 standard; Protein; 792 AA.

XX ABB92160;

XX ABB92160;

DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1371.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB ) BAYER AG.

XX Tietjen K, Weidner M;









GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: June 4, 2003, 14:47:01 ; Search time 16.8268 Seconds  
(without alignments)  
3719.278 Million cell updates/sec

Title: US-09-980-771a-5  
Perfect score: 3370  
Sequence: 1 ALDIVVAAEVAPEWPKTGG.....SASKISAAPLVSAATRKA 651

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PTR-73.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1481.5	44.0	608	2	S43341	starch synthase (E
2	1457.5	43.2	607	1	TOPOX	starch synthase (E
3	1415.5	42.0	603	2	S61504	glycogen(starch) s
4	1412.5	41.9	608	2	T10906	starch synthase (E
5	1398.5	41.5	609	2	JQ0703	glycogen(starch) s
6	1397.5	41.5	609	2	S11481	glycogen(starch) s
7	1395.5	41.4	603	1	YUBHY	glycogen(starch) s
8	1390	41.2	610	2	F86453	granule-bound star
9	1362	40.4	615	1	YUWY	glycogen(starch) s
10	1355.5	40.2	608	2	T14731	glycogen(starch) s
11	1353.5	40.2	605	1	S07314	glycogen(starch) s
12	1194	35.4	238	2	T07921	probable starch sy
13	902.5	25.8	641	2	T07668	starch synthase (E
14	873.5	25.9	626	2	TQ2322	starch synthase (E
15	872	25.9	732	2	T01208	starch synthase (E
16	871	25.8	752	2	S61505	glycogen(starch) s
17	869	25.8	610	2	T06280	probable starch sy
18	861.5	25.6	622	2	T01414	starch synthase (E
19	847.5	25.1	788	2	T07667	glycogen(starch) s
20	839	24.9	491	2	T06798	probable starch sy
21	837	24.8	698	2	T01209	starch synthase (E
22	684.5	20.3	477	2	B95130	glycogen synthase
23	683.5	20.3	477	2	H98000	starch synthase (E
24	677.5	20.1	484	2	S40051	starch synthase (E
25	669	19.9	483	2	C86712	starch synthase (E
26	652	19.3	480	2	H98228	glycogen synthase
27	652	19.3	480	2	AF3057	glycogen synthase
28	651	19.3	486	2	H72321	glycogen synthase
29	637	18.9	477	2	A97176	glycogen synthase,

30 622 18.5 484 2 F82165 glycogen synthase  
31 614.5 18.2 476 2 E83785 starch (bacterial  
32 604.5 17.9 486 2 C96018 probable starch sy  
33 604 17.9 476 2 AE0479 starch synthase (E  
34 590.5 17.5 530 2 AH3194 glycogen synthase  
35 587 17.4 477 1 SYECGL starch synthase (E  
36 587 17.4 477 2 B98163 glycogen synthase  
37 587 17.4 477 2 C86009 glycogen synthase  
38 587 17.4 477 2 AI0995 starch synthase (E  
39 569 16.9 476 2 C64119 starch synthase (E  
40 568.5 16.9 1230 2 T07663 soluble starch syn  
41 561.5 16.7 1025 2 H86250 hypothetical prote  
42 560.5 16.6 472 2 AI2040 glycogen synthase  
43 552.5 16.4 463 2 C70363 glycogen synthase  
44 536 15.9 492 2 AG1810 glycogen (starch)  
45 535 15.9 477 2 S76496 hypothetical prote

## ALIGNMENTS

RESULT 1

S43341

starch synthase (EC 2.4.1.21) precursor - cassava

N:Alternate names: bacterial-glycogen synthase; glycogen synthase; starch synthase  
C:Species: Manihot esculenta (cassava)

C>Date: 07-Sep-1994 #sequence\_revision 03-Nov-1995 #text\_change 16-Aug-2002

C:Accession: S43341

R:Salehuzzaman, S.N.I.M.; Jacobsen, E.; Visser, R.G.F.

Plant Mol. Biol. 23, 947-962, 1993

A:Title: Isolation and characterization of a cDNA encoding granule-bound starch synth

A:Reference number: S43341; MUID:94083565; PMID:8260633

A:Accession: S43341

A:Molecule type: mRNA

A:Residues: 1-608 <SAL>

A:Cross-references: EMBL:X74160; NID:g437041; PIDN:CAA52273.1; PID:g437042

C:Genetics:

A:Gene: GBSS; waxy

A:Genome: nuclear

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produc

C:Superfamily: starch synthase

C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltra

F:1-78/Domain: transit peptide (amyloplast) #status predicted <TNP>

F:79-608/Product: ADPglucose-starch glucosyltransferase #status predicted <MAT>

Query Match 44.0%; Score 1481.5; DB 2; Length 608;

Best Local Similarity 55.0%; Pred. No. 4.4e-81;

Matches 299; Conservative 64; Mismatches 148; Indels 33; Gaps 9;

QY 2 LDIYVAAEVAPEWPKTGGTGLDVTGGLPLEYKRGHRVMTIAPRYDOYADAMDTSVVVDI- 60

Db 82 MNLFLVGAEEVGPWSEKTKGLDGLPDPAMAARGHRVMTVSPRDOYKADMTSVSVSEIK 141

QY 61 MG---EKVYRPHSTKKGVRHVRWIDHPWFLAKWVGKTSKLYGPRSGADYLDNHRKFLFC 117

Db 142 IGDRIETVRFPHSKRGVDRVDFVDPMPLEKVGWGTSGIKYGPAGLDYQDNQLRSLC 201

QY 118 KAAEAAARVLFF-----GP-GEDCVTVANDHSALVPVLLKDEYQPKGFTKAKSVLA 169

Db 202 LAALAEPRVLMNSKNKSGPYGEVAFIANDWHTALLPCYLKAIYQPMGIYKHAQVAF 261

QY 170 IHNIAFGRWEEAFKDTKLPQAAFDKLAFSDGAKYVTEATPMEDEKPELTKTKYKKI 229

Db 262 IHNIAFGRAFSDFPRLNLPDKFSFDFIDGYEK-----PVKG---RKI 304

QY 230 NLWKGIIADKNTVSPNTATETAADAAGVELDTVIRAKGIEGVINGMDIEENPKTD 289

Db 305 NWMKAGILESRLVTSPPYTAQEVISGVERGVELDNIRKTTGIAIINGMDVQWNPVD 364

QY 290 KFLSAPYDQNSVYAGKAAKFAQLGELPVDTPATPFIATIGRLEEGQGVDTIIAALPKI 349

Db 365 KYIDHYDATTVMADKPLLKALQAEGLPVDNRVPLIGFTIGRLEEGQGSDFVAAISQ- 423



Db 197 QAALAPRVNLNSKSFSGYGEDVIFVANDHWSALIPCYLKSWMYKSRGLYKNAKVAFC 256  
 QY 170 IHNIAFOGRMWEBAFKD---TKLPQAAFDKLAFSDGYAKYVTEATPMEDEKPLTGTGTY 226  
 Db 257 IHNIAOGR---NAFSDFLNLNLPDEPRSSDFIDGYNK-----PCEG--- 296  
 QY 227 KKNLWKGGLIADKLVTPSPNATEAADAAGGVELDTVIRAKGIEGIVNGMDIEEWN 286  
 Db 297 KKNLWKGGLIADKLVTPSPNATEAADAAGGVELDTVIRAKGIEGIVNGMDIEEWN 356  
 QY 287 KTKFSLAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEGKGVDIILAAAL 346  
 Db 357 QTDRIYDHYNEVTEAKPLKCTLOAEGLPVDSSIPILGIFGRLEEGKGVDIILVFAI 416  
 QY 347 PKTILATPKVOIALGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFM 406  
 Db 417 AK-FADENVOIVLGTGKTKMEQKQIELEEKYKPKAIGITKFNPSPLAHKIAGADFIVP 475  
 QY 407 SRPEPCGLIOLHAMHYGTVPVASTGGVLDVTKGVTGFHMGALNP--DKLDEADADALA 464  
 Db 476 SRPEPCGLIOLHAMHYGTVPVASTGGVLDVTKGVTGFHMGALNP--DKLDEADADALA 535  
 QY 465 ATVRASEVFAGGRYPENVANCISQDLSWSKPAOKWEGLLFEV-----VYKGG--VATA 517  
 Db 536 ATKRALKTYGTQAMKQIILNCMAQNFSSWKKPAKLWEKALLNLEVTGNVAGIDGDEIAPL 595  
 QY 518 KKEIKVP 525  
 Db 596 AKENVATP 603

## RESULT 4

T10906

starch synthase (EC 2.4.1.21) - sweet potato

N:Alternate names: starch synthase

C:Species: Ipomoea batatas (sweet potato)

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Aug-2002

C:Accession: T10906

R:Wang, S.J.; Yeh, K.W.; Tsai, C.Y.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z17212

A:Accession: T10906

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-608 &lt;WAN&gt;

A:Cross-references: EMBL:U44126; NID:g1172158; PID:g1172159

A:Experimental source: cv. Tainong; tuberous root

C:Genetics:

A:Gene: SS67

C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing a

A:Pathway: starch biosynthesis

C:Superfamily: starch synthase

C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 41.98; Score 1412.5; DB 2; Length 608;

Best Local Similarity 53.58; Pred. No. 66/77;

Matches 292; Conservative 61; Mismatches 152; Indels 41; Gaps 10;

QY 2 LDIWVAEAVAPWSKSTGGGLGVTGGLPIELVKRGHVRVMTIAPRYDQYADAWDTSVVVD-I 60

Db 82 MNLVFCGEGPCKTGGGLGVTGGLPIELVKRGHVRVMTIAPRYDQYADAWDTSVVVD-I 141

QY 61 MG---EKVRVPHSKGVRHWIDHPWFLAKVWCKTSGKLYGPRSGADYLDNHRKRFALFC 117

Db 142 VGDRIEVRPHSKGVRHWIDHPWFLAKVWCKTSGKLYGPRSGADYLDNHRKRFALFC 201

QY 118 KAALAEARVLPF-----GP-GEDCVFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 169

Db 202 QAALAEARVNLNSKSFSGYGEDVIFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 261

QY 170 IHNIAFOGRMWEBAFKDTKLPQAAFDKLAFSDGYAKYVTEATPMEDEKPLTGTGTYKKI 229

Db 262 IHNIAOGRFAFSDFLNLNLPDEYKGSFDFIDGTDK-----PVKG---RKI 304  
 QY 230 NNLKGGIITADKLVTPSPNATEAADAAGGVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289  
 Db 305 NNMKAGIREADRVFTSPNYAKELVSCVSKGVDELNDHIRDCGIGTCGMDTQEWNPATD 364  
 QY 290 KFLSAPYDQNSVYAGKAAAEALQAEGLPVDPTAPLFAFGRLEEGKGVDIILAAALPKI 349  
 Db 365 KYLAVKDYITITVMAQKPLLEALQAAVGLPVDRIPLTIGFGRLEEGKGSILYAAISKF 424  
 QY 350 LATPKVOIALGTGKAAEYKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRF 409  
 Db 425 ISM-DVQILLIGTGGKFEQIEQLEVMYPDKRGVAKFNVPPLAHMITAGADFMVPSRF 483  
 QY 410 EPCGLIOLHAMHYGTVPVASTGGVLDVTKGVTGFHMGALNP--DKLDEADADALAATV 467  
 Db 484 EPCGLIOLHAMRYGTPCICASTGGVLDVTKGVTGFHMGALNP--DKLDEADADALAATV 543  
 QY 468 RASEVFAGGRYPENVANCISQDLSWSKPAOKWEGLL-----EEVYKGGVATAK 519  
 Db 544 GRALITGTGTAFTETIMIKNCSQELSWKGPKNWETVLLSLGVAGSEPGVEGE-ETIAPLAK 602  
 QY 520 BEIKVP 525  
 Db 603 ENVATP 608

RESULT 5  
 JQ0703  
 glycogen(starch) synthase (EC 2.4.1.11) - rice  
 N:Alternate names: granule-bound starch synthase  
 C:Species: Oryza sativa (rice)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Aug-2002  
 C:Accession: JQ0703  
 R:Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.  
 Nucleic Acids Res. 18, 5898, 1990  
 A:Title: Nucleotide sequence of rice waxy gene.  
 A:Reference number: JQ0703; MUID:91016948; PMID:2216792  
 A:Accession: JQ0703  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-609 <WAN>  
 A:Cross-references: EMBL:X53694  
 A:Experimental source: strain subsp. japonica Hengfeng  
 C:Genetics:  
 A:Gene: waxy  
 A:Introns: 114/1; 174/1; 204/1; 225/2; 259/1; 295/3; 377/1; 436/1; 500/1; 529/1  
 C:Function:  
 A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose produci  
 C:Superfamily: starch synthase  
 C:Keywords: chloroplast; glycogen/starch biosynthesis; glycosyltransferase; hexosylt

Query Match 41.58; Score 1398.5; DB 2; Length 609;  
 Best Local Similarity 51.28; Pred. No. 4.1e-76;  
 Matches 280; Conservative 77; Mismatches 147; Indels 43; Gaps 9;  
 QY 2 LDIWVAEAVAPWSKSTGGGLGVTGGLPIELVKRGHVRVMTIAPRYDQYADAWDTSVVVD-I 60  
 Db 83 MNLVFCGEGPCKTGGGLGVTGGLPIELVKRGHVRVMTIAPRYDQYADAWDTSVVVD-I 142  
 QY 61 ---MGEKVRVPHSKGVRHWIDHPWFLAKVWCKTSGKLYGPRSGADYLDNHRKRFALFC 117  
 Db 143 VADRIEVRPHSKGVRHWIDHPWFLAKVWCKTSGKLYGPRSGADYLDNHRKRFALFC 202  
 QY 118 KAALAEARVLPF-----PFGP---GEDCVFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 169  
 Db 203 QAALAEARVNLNNPFPKTYGDEVFVANDHWSALVPVLLKDEYQPKGFTKAKSVLA 262  
 QY 170 IHNIAFOGRMWEBAFKDTKLPQAAFDKLAFSDGYAKYVTEATPMEDEKPLTGTGTYKKI 229  
 Db 263 IHNISYOCGRFAFEDYPELNLSEFRSSDFIDGY-----DTPVEG-----RKI 305  
 QY 230 NNLKGGIITADKLVTPSPNATEAADAAGGVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289

Db	306	NMKAGILEADRVLTSPYYAEELISGTARGCELDNIMRTGTITGVNGDVMSEWPSDK	365
QY	290	KFLSAPYQDQNSVYAGKAAAKALQAEGLPVDPTAPLFAFIRGLEQKGVDITLAAIPKI	349
Db	366	KYITAKYDATTAEAKALNKEALQAEAGLPVDRKIPLIAFIRGLEQKGPDVMAAIPEL	425
QY	350	LATPKVQIALLTGKAAAYEKLIVNAIGTKYKGRAGVVKFSAPLAHMLTAGADEFMLVPSRF	409
Db	426	M-QEDVQIQLVLLTGKKKPEKLLKSWEKYPGKRVAVVKFNAPLAHLMAGADVLVPSRF	484
QY	410	EPGGLIQLHAMHYGTVPVASTGGLVDTVVEGYTGFFHMGALNPD--KLDEADADALAATV	467
Db	485	EPGGLIQOGMRGYGTACACASTGGLVDTVTEGKTGFHMRGLSDVGKWFPSDVKKVAATL	544
QY	468	RAASEVFAAGGRYPENVANCISQDLSWSKPAQKEGGLEEVVYCKGG-----VATAK	518
Db	545	KRAIKVGTGPAYEEWRNCMDQLSWGPKAKNNENVL--LGLGVAGSAPGIEGDEIRPLA	602
QY	519	KEEIKVP	525
Db	603	KENVAAP	609

## RESULT 6

SL1481  
glycogen (starch) synthase (EC 2.4.1.11) precursor - rice  
N;Alternate names: starch synthase; waxy protein  
C;Species: *Oryza sativa* (rice)  
C;Date: 05-Mar-1995 #sequence.revision 10-Nov-1995 #text\_change 16-Aug-2002  
C;Accession: S11481; S22519; S30485; PC2190; JQ2224  
R;Wang, Z.Y.; Wu, Z.L.; Xing, Y.Y.; Zheng, F.G.; Guo, X.L.; Zhang, W.G.; Hong, M.M.  
Nucleic Acids Res. 18, 5898, 1990  
A;Title: Nucleotide sequence of rice waxy gene.  
A;Reference number: JQ0703; MUID:91016948; PMID:2216792  
A;Accession: S11481  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-609 <WAX>  
A;Cross-references: EMBL:X53694; NID:g577598; PIDN:CAA37732.1; PID:g577599  
R;Okagaki, R.J.  
Plant Mol. Biol. 19, 513-516, 1992  
A;Title: Nucleotide sequence of a long cDNA from the rice waxy gene.  
A;Reference number: S22519; MUID:92322985; PMID:1377969  
A;Accession: S22519  
A;Molecule type: mRNA  
A;Residues: 1-609 <OKA>  
A;Cross-references: EMBL:X62134; NID:g20402; PIDN:CAA44065.1; PID:g20403  
R;Hirano, H.Y.; Sano, Y.  
Plant Cell Physiol. 32, 989-997, 1991  
A;Title: Molecular characterization of the waxy locus of rice (*Oryza sativa*).  
A;Reference number: S30485  
A;Accession: S30485  
A;Status: preliminary; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-609 <HIR>  
A;Cross-references: EMBL:X58228; NID:g20400; PIDN:CAA41186.1; PID:g20401  
A;Accession: PC2190  
A;Molecule type: protein  
A;Residues: 78-113 <HIR>  
A;Experimental source: leaf, cDNA POSLHC2120  
C;Comment: This protein is involved in amylose synthesis in the rice endosperm.  
C;Genetics:  
A;Gene: waxy  
A;Introns: 113/3; 140/3; 173/3; 203/3; 225/1; 258/3; 295/2; 376/3; 435/3; 499/3; 528/3;  
C;Function:  
A;Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose producing  
A;Pathway: starch biosynthesis  
C;Superfamily: starch synthase  
C;Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltransferase;  
F;1-77/Domain: transit peptide (amyloplast) #status predicted <TNP>  
F;78-609/Product: glycogen (starch) #status experimental <NAT>



[illegible]

```

RESULT 11
S07314
glycogen(starch) synthase (EC 2.4.1.11) precursor - maize
N:Alternate names: starch synthase; UDP-glucose starch glycosyltransferase, sta
C:Species: Zea mays (maize)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Aug-2002
C:Accession: S07314
R:Kloesgen, R.B.; Gierl, A.; Schwarz-Sommer, Z.; Saedler, H.
Mol. Gen. Genet. 203, 237-244, 1986
A:Title: Molecular analysis of the waxy locus of Zea mays.
A:Reference number: S07314
A:Accession: S07314
A:Molecule type: DNA
A:Residues: 1-605 <KLO>
A:Cross-references: EMBL:X03935; NID:g22509; PIDN:CAA27574.1; PID:g1644339
A:Experimental source: line C
A:Title: translation of the nucleotide sequence is not complete
A:Note: part of this sequence, including the amino end of the mature protein, w
C:Genetics:
A:Gene: waxy
A:Introns: 107/3; 134/3; 167/3; 197/3; 219/1; 252/3; 289/2; 370/3; 431/3; 495/3
C:Function:
A:Description: catalyzes the alpha-1,4-glucosylation of starch by UDPglucose pr
C:Superfamily: starch synthase
C:Keywords: amyloplast; chloroplast; glycogen/starch biosynthesis; glycosyltra
F:1-72/Domain: transit peptide (amyloplast) #status predicted <NP>
F:73-605/Product: glycogen (starch) synthase #status experimental <MAT>

Query Match          40.2%  Score 1353.5;  DB 1;  Length 605;
Best Local Similarity 50.4%  Pred. No. 2e-73;
Matches 275;  Conservative 79;  Mismatches 153;  Indels 39;  Gaps 1

```





Db 588 ---LUTLKLAI-----GTYTEHKSSWEGLMRRMGDRYSWENAAIQYE 627

## RESULT 14

starch synthase (EC 2.4.1.21) precursor - rice  
N:Alternate names: starch synthase  
C:Species: Oryza sativa (rice)  
C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 16-Aug-2002  
C:Accession: JQ2322; PQ0811  
R:Baba, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shimada, H.; Kobayashi, E.; Ohnishi  
Plant Physiol. 103, 565-573, 1993  
A:Title: Identification, cDNA cloning, and gene expression of soluble starch synthase in  
A:Reference number: JQ2322; MUID:94302151; PMID:7518089  
A:Accession: JQ2322

A:Molecule type: mRNA  
A:Residues: 1-626 <BAB1>  
A:Cross-references: DBJ:16202; NID:9450484; PIDN:BAA03739.1; PID:9450485  
A:Accession: PQ0811  
A:Molecule type: protein  
A:Residues: 114-129 <BAB2>  
A:Experimental source: seed  
C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose producing  
C:Superfamily: starch synthase  
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase  
F:1-113/Domain: transit peptide (amyloplast) #status predicted <TRS>  
F:114-626/Product: ADPglucose-starch glucosyltransferase, 57KD form #status experimental  
F:122-626/Product: ADPglucose-starch glucosyltransferase, 55KD form #status experimental  
F:440-448/Region: substrate binding #status predicted

Query Match 25.9%; Score 873.5; DB 2; Length 626;  
Best Local Similarity 41.5%; Pred. No. 1.1e-44;  
Matches 197; Conservative 77; Mismatches 154; Indels 47; Gaps 13;

QY 4 IVVAAAEVAPWSTKGLGDTGGLPIELVKRHRVMTIAPRYDQYADAWDTSV 57

Db 135 VVFTGEASPYAKSGGLGDCVCSLPTALALRGHRVWVMPRYMNGALNKNFANAFYTEKH 194

QY 58 VDTM---GE-KVRYHSIKGVHRVWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNKR 113

Db 195 IKPCFGEHEVTFEYRSDVDWVDFHPSY-----HRPNLYGDNFCA-FGDNQF 247

QY 114 ALFCAAEAEARVLPGP---GEDCVFVANDHISALVPVLLKDEYQPKGFTTRAKSVLA 170

Db 248 TLCLYAAACEAPLLELGGYTYGCKMVFVNDHWSALVPLVLAARPYGVYRDARSVLVI 307

QY 171 HNIAGQRMWEAEAFKDTKLPAQAFDLAFSDGYAKVYVTEATPMEDEKPLTGKTYKIN 230

Db 308 HNLHOGVEPASTYDGLPPEWYGALW-----VPEWARHLDKG-----EAVN 354

QY 231 WLKGGIADKLVTSPNATEIAADAAGVELDVI--RAKGIEGIVNGMDEEWNPK 288

Db 355 FLKGAVTADRIVTVSQGSWEVTT-AEGGGLNELLSSKSVLNGIVNGIDINDNPST 413

QY 289 DKFLSAPDQNSVYAKAAKALQAEGLPVDPTAPLFAFGRLEEQKGVDIILALPK 348

Db 414 DKFLPYHSVDL-SGAKCAKAEQLGELPIRPDVLPIGFLIGRLDYGIDILKLAIPD 472

QY 349 ILATPKVQIATLGTKAAVEKLVNAIGTKYKRAKGVVYKFSAPLAHMLTAGADFMVPSR 408

Db 473 LM-RDNIQVLMGSDGDFEGHWRSTESYRDKFRGWGFSVPVSHRITAGDILMPSR 531

QY 409 FPFCGLIQLHAMHYGTVPVVASTGGLVDTVK-----EGVTGFHMGALNPK 454

Db 532 FEPCLNQLYAMQYGVVPHVHTGGLRDTVENFNFPAEKGEQGTGWAFTPLIEK 586

## RESULT 15

T01208

starch synthase (EC 2.4.1.21) isoform STSII-1 - maize (fragment)

N:Alternate names: starch synthase isoform STSII-1

C:Species: Zea mays (maize)  
C>Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Aug-2002  
C:Accession: T01208  
R:Knight, M.E.; Harn, C.; Lilley, C.E.R.; Guan, H.P.; Singletary, G.W.; Mu-Forester,  
Plant J. 14, 613-622, 1998  
A:Title: Molecular cloning of starch synthase I from maize (W64A) endosperm and expr  
A:Reference number: Z14279; MUID:98340555; PMID:9675904  
A:Accession: T01208

A:Status: translated from CB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-732 <KN1>  
A:Cross-references: EMBL:AF019296; NID:92811133; PIDN:AAD13341.1; PID:g2655029  
A:Experimental source: strain W64A; endosperm  
C:Genetics:  
A:Gene: SSIIa  
C:Function:

A:Description: catalyzes the alpha-1,4-glucosylation of starch by ADPglucose produci  
C:Superfamily: starch synthase  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.9%; Score 872; DB 2; Length 732;  
Best Local Similarity 39.7%; Pred. No. 1.7e-44;  
Matches 211; Conservative 77; Mismatches 178; Indels 66; Gaps 15;

QY 2 LDIVAAAEVAPWSTKGLGDTGGLPIELVKRHRVMTIAPRYDQYADAWDTSV-VVD 59

Db 241 MNVIVAAECPWCKTGLGSDVGVGALPKALARRHRVWVVPVRYGDIYVEAFDMGIRKYYK 300

QY 60 IMGE-KVRYHSTKGVHRVWIDHPWFLAK---VWGKTGSKLYGPRSGADYLDNKRHKA 114

Db 301 AAGDQLEVNFHAFIDGVDFVDAPLFRHQDDIYGSRQFIM-----KRM 348

QY 115 LFCAAEAEARVLP-----FGEDCVFVANDHISALVPVLLKDEYQPKGFTTRAKSVLA 169

Db 349 LFCVAVVEVPHVPCGGVCGYGDG-NLVFIANDHWTALLPYLYKAYIRHGLMQYTRSLV 407

QY 170 IHNIAFGQRMWEAEAFKDTKLPAQAFDLAFSDGYAKVYVTEATPMEDEKPLTGKTYK 229

Db 408 IHNIAHOGVPEVDFPYMDLPEHYLQHFELD-----PVGG---EHA 446

QY 230 NLKGGIADKLVTSPNATEIAADAAGVELDVI--RAKGIEGIVNGMDEEWNPK 287

Db 447 NIFAAGLKMADRVVTVSRGYLWELKT-VEGGWGLHDIIRSNWDKINGIVNGIDHQQWNP 505

QY 288 TDKFLSAP---YDQNSVYAKAAKALQAEGLPVDPTAPLFAFGRLEEQKGVDIIL 343

Db 506 VDHLRSDGYTNSLETLDAGKROCKAALQRELGLGVDRDDVPLLGLIGRLDGKGVDI 565

QY 344 AALPKIATPKVQIATLGTKAAVEKLVNAIGTKYKRAKGVVYKFSAPLAHMLTAGADFM 403

Db 566 DAMPWI-AGODVQLVMLGTGGRADLERMLQHLREHPNKGWGVGFSVPMARITAGADVL 624

QY 404 LVPSRPEPCGLIQLHAMHYGTVPVVASTGGLVDTVK-----GVTFGHMGALNPKLDEAD 459

Db 625 VMSRPEPCGLNQLYAMAYGVVPHVHAGVGLRDTVAPDFDFGDLGW-----TFDRAE 678

QY 460 ADALAAVTRRASVEA--GGYRPEMVAANCISQDLSWSKPAQKNEGGLLEVVY 509

Db 679 ANKLIEALRHCLDITYRKYGESWKSQARGMSQDLSWDHAAELYEDVLVAKY 730

Search completed: June 4, 2003, 14:57:31  
Job time : 21.8268 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:50:41 ; Search time 8.74119 Seconds  
(without alignments)  
3088.950 Million cell updates/sec

Title: US-09-980-771a-5

Perfect score: 3370

Sequence: 1 ALDIVNVAEAPWSKTGL.....SASKTSAKPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481.5	44.0	608	1	UGST_MANES
2	1457.5	43.2	607	1	UGST_SOLTU
3	1455.5	43.2	608	1	UGST_ANTMA
4	1415.5	42.0	603	1	UGST_PEA
5	1412.5	41.9	608	1	UGST_IPOBA
6	1397.5	41.5	609	1	UGST_ORYSA
7	1395.5	41.4	603	1	UGST_HORVU
8	1394.5	41.4	609	1	UGST_ORYGL
9	1362	40.4	615	1	UGST_WHEAT
10	1355.5	40.2	608	1	UGST_SORBI
11	1353.5	40.2	605	1	UGST_MAIZE
12	902.5	26.8	641	1	UGS2_SOLTU
13	873.5	25.9	626	1	UGS2_ORYSA
14	871	25.8	752	1	UGS3_PEA
15	869	25.8	610	1	UGS2_WHEAT
16	847.5	25.1	788	1	UGS3_SOLTU
17	684.5	20.3	477	1	GLGA_STRPN
18	677.5	20.1	484	1	GLGA_BACSU
19	670.5	19.9	485	1	GLGA_BACST
20	669	19.9	478	1	GLGA_LACLA
21	653.5	19.4	480	1	GLGL_RHIME
22	652	19.3	480	1	GLGA_AGRIT5
23	651	19.3	486	1	GLGA_THEMEA
24	646.5	19.2	480	1	GLGA_RHITR
25	637	18.9	477	1	GLGA_CLOAB
26	622	18.5	482	1	GLGA_CLOPE
27	622	18.5	484	1	GLGA_VIBCH
28	614.5	18.2	476	1	GLGA_BACHD
29	604.5	17.9	486	1	GLG2_RHIME
30	604	17.9	476	1	GLGA_YERPE
31	602.5	17.9	481	1	GLGA_RHULO
32	587	17.4	477	1	GLGA_ECOLI
33	587	17.4	477	1	GLGA_SALT1

34 580 17.2 477 1 GLGA\_SALT1  
35 578 17.2 480 1 GLGA\_PASMU  
36 569 16.9 476 1 GLGA\_HAEIN  
37 568.5 16.9 1230 1 UGS4\_SOLTU  
38 560.5 16.6 472 1 GLGA\_ANASP  
39 555 16.5 461 1 GLGA\_FUSNN  
40 552.5 16.4 463 1 GLGA\_AQUAE  
41 538 16.0 465 1 GLGA\_SYNP7  
42 536 15.9 477 1 GLGA\_RHOSH  
43 536 15.9 492 1 GLG2\_ANYP3  
44 535 15.9 477 1 GLGA\_SYNP3  
45 527 15.6 444 1 GLGA\_DEIRA

#### ALIGNMENTS

##### RESULT 1

UGST\_MANES STANDARD; PRT; 608 AA.  
AC Q43784;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Granule-bound glycogen [starch] synthase, chloroplast precursor  
(EC 2.4.1.11).  
DE WAXY OR GBSS.  
GN Manihot esculenta (Cassava) (Manioc).  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Manihot.  
OX NCBI\_TaxID=3983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. M.COL 22; TISSUE=Tuberous root;  
RX MEDLINE=94083565; PubMed=8260633;  
RA Salehuzzaman S., Jacobsen E., Visser R.G.F.;  
RT "Isolation and characterization of a cDNA encoding granule-bound  
starch synthase in cassava (Manihot esculenta Crantz) and its  
antisense expression in potato."  
PL Plant Mol. Biol. 23:947-962(1993).  
RL  
CC -!- FUNCTION: RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE IN RESERVE  
STARCH  
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =  
UDP + [(1,4)-alpha-D-glucosyl](N+1).  
CC -!- PATHWAY: Starch biosynthesis.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN A NUMBER OF DIFFERENT ORGANS,  
BUT MOST ABUNDANTLY IN TUBERS.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE  
FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X74160; CAA52273.1; .  
CC InterPro: IPR001296; Glycos\_transf\_1.  
CC Pfam: PF00534; Glycos\_transf\_1; 1.  
CC Glycogen biosynthesis; transferase; Glycosyltransferase;  
KW Glycogen biosynthesis; transferase; Starch biosynthesis.  
KW Transit peptide; Chloroplast; Starch biosynthesis.  
FT TRANSIT 1 78 CHLOROPLAST [BY SIMILARITY].  
FT CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.  
FT BLINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).  
SQ SEQUENCE 608 AA; 66968 MW; C9C970CD3011BDBB CRC64;

Query Match 44.0%; Score 1481.5; DB 1; Length 608;  
Best Local Similarity 55.0%; Pred. No. 2.4e-81;  
Matches 299; Conservative 64; Mismatches 148; Indels 33; Caps 9;

Qy	2	LDIVMAAEVAPSWK	TGGIGDVTGG	PLTELVKGRHVRMT	IAPRYDOYADAWDT	SVVVDI-	60
		:::					
Db	82	NNLIPVGA	EVGPWSKT	GGLGDLVGL	GPMAARGHRVMT	VSRYDOYKAWDT	SVSVEIK 141
		:::					
Qy	61	MG----	EKVRYHSH	IKKGVRHVR	WHDWFLAK	YWGKTSGL	KYGRSGADYLDNHKRALFC 117
		:					
Db	142	IGDRIETV	RRFHSYKRG	VDVDFVDP	MELEKVMG	KTSKIYGR	AGLDYQDNQLRSLIC 201
		:					
Qy	118	KAATEAAR	VLFP-----	GP-GE	DCVAVAND	HSALVP	VLKDBYQPGQFTKAKSVLA 169
		:					
Db	202	LAALAPR	VLNLNSK	NFSGPGY	EEVAFIAND	WHTALT	PCYLKAIYPMGIYKHAKVAFC 261
		:					
Qy	170	IHNIAFOGR	MWEEAFKDT	KLQAA	DFDKLAF	SDGYAKV	TEATPMEEDEKPLTKGYTKKI 229
		:					
Db	262	IHNIAYG	GRFAFSDF	PRNLN	PKFKSS	DFIDG	EYK-----PVKG---RKI 304
		:					
Qy	230	NWLKGGI	IAADKLVT	VSNYATE	IAADA	AGGVGL	DTVIRAKGIEGVANGMTEENPKTD 289
		:					
Db	305	NWMAKGL	LES	DRVLT	VSPIYA	QAEVTS	GVGVGLDNFRKTGTGIANGMDVQENPVD 364
		:					
Qy	290	KFLSAPY	DONS	VYAGKAA	KALQAE	LGLP	VDPTAPLFAFGRLEBQGVDIILAALPKI 349
		:					
Db	365	KYDIDH	DATVYMD	AKP	LLKALQAE	VLG	VDNRNVLPTGFTGRLEBQGSDFVAAISO- 423
		:					
Qy	350	LATPKVQ	IA	LGTG	KAAYEK	KLIVNA	IGTKYKGRAGVVKFSAPLAHMLTAGADFMLVPGRF 409
		:					
Db	424	LVEHN	VQIVIL	GTGKK	FKFQIE	HLVLY	PDKRGVAKFNVP LAHMITAGADFMLVPGRF 483
		:					
Qy	410	EPCLLIQ	LHAMH	GTVP	VAVAST	TGGLV	DTVKEGVTGFHMGALNP--DKLDEADADALAATV 467
		:					
Db	484	EPCLLIQ	LHAMR	GTVP	IVAST	GTGLV	DTVKEGVTGFQMGALUHVCECDKIDSADVAIAIVKTV 543
		:					
Qy	468	RRASEV	FAGGRY	PEMWAN	TSODLS	WSKPAQ	KWESGLLEEYVYGGGVATAKKEEIKVPVA 527
		:					
Db	544	ARALGY	ATAAL	REMILN	CAQDLS	WSKPAR	WEKMLLDLEV-TGSEFGTGEETAPLAK 602
		:					
Qy	528	EKTP	531				
		:					
Db	603	ENVP	606				
		:					

## RESULT 2

UGST_SOLUTU	STANDARD;	PRT;	607 AA.
ID	UGST_SOLUTU		
AC	Q00775; Q43176;		
AD	Q00775; Q43176;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Granule-bound glycogen [starch] synthase, chloroplast precursor		
DE	(EC 2.4.1.11).		
DE	WAXY OR GBSS.		
OS	Solanum tuberosum (Potato).		
OS	Solanum tuberosum (Potato).		
OC	Eukaryota; Viridiplantae;		
OC	Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
OC	Assteridae; euasterids I; Solanales; Solanaceae; Solanum.		
OX	NCBI_TaxID=4113;		
OX	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RP	STRAIN=AM79.7322;		
RRX	MEDLINE=91360072; PubMed=1886609;		
RRX	van der Leij F.R., Visser R.G.F., Ponstein A.S., Jacobsen E.,		
RRR	Feenstra W.J.;		
RRR	"Sequence of the structural gene for granule-bound starch synthase		
RRRT	potato (Solanum tuberosum L.) and evidence for a single point		
RRRT	deletion in the amf allele."		
RRRT	Mol. Gen. Genet. 228:240-248(1991).		
RRRN	[2]		
RRP	REVIEWS.		
RRP	van der Leij F.R.;		
RRR	Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.		
RRRL	[3]		
RRRN	SEQUENCE FROM N.A.		
RRRN	RRP		

```

CC STRAIN=cv. Dongnong 303;
CC Dai W.L., Deng W., Cui M., Xiu M., Zhao S.Y., Wang X.M.;
CC RL Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
CC     UDP + [(1,4)-alpha-D-glucosyl](N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC     FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58453; CAA41359.1; -.
CC EMBL; X83220; CAA58220.1; -.
CC PIR; SL655; YUPOY.
CC InterPro: IPR001296; Glycos_transf_1.
CC Pfam; PF00534; Glycos_transf_1; 1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT      1 77
CC              CHLOROPLAST.
CC CHAIN        78   607
CC BINDING      95   95
CC CONFLICT    130  130
CC CONFLICT    398  398
CC SEQUENCE    607 AA; 66575 MW; 2A377865CFFAF650 CRC64;
CC FT
CC FT
CC FT
CC FT
CC FT
CC FT

```

### Query Match

[illegible]

```

Db      602 ENVATP 607

RESULT 3
UGST_ANTMA STANDARD; PRI; 608 AA.
AC O82627;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11) (GBSS1).
GN WAXY OR GBSS.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX PubMed=10364391;
RA Merida A., Rodriguez-Galan J.M., Vincent C., Romero J.M.;
RT "Expression of the granule-bound starch synthase 1 (Waxy) gene from
RT snapdragon is developmentally and circadian clock regulated.";
RL Plant Physiol. 120:401-410(1999).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- TISSUE SPECIFICITY: In leaves, flowers and fruits. Observed in all
CC floral whorls at early developmental stages, but restricted to
CC carpel before anthesis.
CC -!- INDUCTION: Expressed with a circadian rhythm with peak expression
CC at the end of the day.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ006293; CAA06958.1; -
CC EMBL; AJ006294; CAA06959.1; -
CC InterPro; IPR001296; Glycos_transf_1.
CC Pfam; PF00534; Glycos_transf_1; 1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 79 608 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
CC BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 608 AA; 66361 MW; 6527D53D56B6E0C CRC64;

Query Match 43.2%; Score 1455.5; DB 1; Length 608;
Best Local Similarity 54.2%; Pred. No. 8.7e-80;
Matches 296; Conservative 61; Mismatches 148; Indels 41; Gaps 10;

QY 2 LDVWVAERVAEPWSKTGGGLDVTGGLPTELVRKGRHVMTIAPRYDQYADAMDTSVVVDI- 60
Db 82 MNLVFLAEVGPWSKTGGGLDVTGGLPAPAMAGNGHRVMTVSPRYDQYKADWTSVVVEIK 141
QY 61 MG---EKVRYFSIKGHRVWIDHFWFLAKVWGKTGSKLYGPRSGADYLDNHHKFPALFC 117
Db 142 VGSDETIVRFPHCYKRGVDRVFEVDHPIFLEKVGKTKSKIYGNAGTDYQDNQLRSLJLC 201
QY 118 KAAITAARVLPF-----GP-GEUCVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169
Db 202 QAALEAPRVNLNTSSKYFSGPYGEDYVVFANDWHTALLPCYLKSMYQSKGMVILHAKVAF 261

```

```

QY 170 IHNIAFOGRWEEAFKDTKLPOAAFDKLAFCGAKYKVTYATPEEEDKPLTKYTKKI 229
Db 262 IHNIAFGGRGSDSCFLNLPDQFKSSEDFDGYEK-----PVKG---RKI 304
QY 230 NMLKGGIITADKLIVTSPNVATETIAADAAGGVBLDTVIRAKGIEGIVNGMDIEWNPKT 289
Db 305 NNMKAGILESDRVVTVPSPYAMELVSGAEKGVLDNVIAKTSITGIVNGMDTQWNPATD 364
QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFIRLEBQKGVDDIILALPKI 349
Db 365 KHTDNYDIITVMDAKPPLLKEALQAAVPLPDKNIPVIGFIRLEEQKGSIIIVAAISKF 424
QY 350 LAIPKVOIALLGTCGAAYEKLVNAIGTKYGRAGKGVVKKFSAPLAHMLTAGADFMVPSRF 409
Db 425 VGL-DVQIIILGTGKKFEQIQIEVLVYDPKARGVAKFNVP LAHMLTAGADFMVPSRF 483
QY 410 EPCGLIQHAMHYGTVPVAVASTGIVDTVKEGVTGFHMGALNP--KLDADADALAATV 467
Db 484 EPCGLIQHAMRYGTIPICASTGGLVDTVTEGTFGHMGAFNVCAITVDPADVOKIATTV 543
QY 468 RRASEVFAGGRYPWEMVANCISQDLWSKPAQKNEGILL-----EEVYKGGGVATKK 519
Db 544 ERALAAYSVAIKEMIQNCMAQDLWSKPAKNWEKMLLSLVGSGSEPGVDGE-ETAPLAK 602
QY 520 EEIKVP 525
Db 603 ENVATP 608

RESULT 4
UGST_PEA STANDARD; PRI; 603 AA.
AC Q43092;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11) (GBSS1).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 76-88.
RC STRAIN=cv. BCI/RR; TISSUE=Embryo;
RX MEDLINE=93251108; PubMed=1302049;
RA Dry I., Smith A., Edwards A., Bhattacharyya B., Dunn P., Martin C.;
RT "Characterization of cDNAs encoding two isoforms of granule-bound
RT starch synthase which show differential expression in developing
RT storage organs of pea and potato.";
RL Plant J. 2:193-202(1992).
CC -!- FUNCTION: MAY BE RESPONSIBLE FOR THE SYNTHESIS OF AMYLOSE.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) =
CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN PODS AND LEAVES. NO EXPRESSION
CC IN FLOWERS OR STIPULES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF EMBRYONIC
CC DEVELOPMENT WITH HIGHEST LEVELS IN LATER DEVELOPMENTAL STAGES.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC

```

```

DR EMBL; X88789; CAA61268.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; Transferase; glycosyltransferase;
KW Transit peptide; Chloroplast; starch biosynthesis.
FT TRANSIT 1 75 CHLOROPLAST.
FT CHAIN 76 603 GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE.
FT BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 603 AA; 63662 MW; 817252FDD12CCAA0 CRC64;

Query Match 42.0%; Score 1415.5; DB 1; Length 603;
Best Local Similarity 51.8%; Pred. No. 2.1e-77;
Matches 284; Conservative 79; Mismatches 140; Indels 45; Gaps 12;

QY 2 LDIVYAAEYAPWPKTGGLGVTGGLPTELVKRGHRVMTIAPRYDQYADAWDTSVVVDI- 60
DB 77 MSLVFVGAEGPWSKTGGLGVLGGLPPVLAGNHRVMTSPRYDQYKDAWDTNVLVEVK 136
QY 61 MGEK---VRYFHSIKKGVRWIDHPWFELAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117
DB 137 VGDRIETVRPHCYKRGVDRVDFVDPHLEFLEKRWGKTGSKLYGPKTGIDYRDNQLRFSLLC 196
QY 118 KAATEARVLPF-----GP-GEDCVFVANDHNSALVPVLLKDEYQPKGFTAKSVLA 169
DB 197 QAALAPRVNLNLSNKKYFSGPYGEDVIFVANDHNSALIPCVLKSMYKSRGLYKNAKVAFC 256
QY 170 IHNIAFGRWEEAFKDTKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTGTGY 226
DB 257 IHNIAFGRWEEAFKDTKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTGTGY 226
QY 227 KKNLWKGIIAADKLVTSPNYATEIADAAGGVLDTVIRAKGIEGVNGMDIEWNP 286
DB 297 KKNWKGIIAADKLVTSPNYATEIADAAGGVLDTVIRAKGIEGVNGMDIEWNP 286
QY 287 KTDKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFTIGRLEQGVDIILAAL 346
DB 357 QTDRIETVRPHCYKRGVDRVDFVDPHLEFLEKRWGKTGSKLYGPKTGIDYRDNQLRFSLLC 416
QY 347 PKILANPKVQIALIGTGKAAEKLIVNAIGTKYGRAGVYVFSAPLAHMLTAGADFMVLP 406
DB 417 AK-FAENVOIVLVGKTKTKMKEQIEVLEKYPGKAIGITKFNPLAKHIIAGADFVIP 475
QY 407 SREPCGLIQAHAMHYGTVPVASTGGLVDTVKRGVTFGRHMGALNP--DKLDEADADALA 464
DB 476 SREPCGLIQAHAMHYGTVPVASTGGLVDTVKRGVTFGRHMGALNP--DKLDEADADALA 464
QY 465 AIVRRASEVYAGRYPEMVANCISQDLSWSKPAQKWEGLLEEV-----VYKGGG--VATA 517
DB 536 AIVRRASEVYAGRYPEMVANCISQDLSWSKPAQKWEGLLEEV-----VYKGGG--VATA 517
QY 518 KKEEIKVP 525
DB 596 AKENVATP 603

RESULT 5
UGST_IPOBA
ID UGST_IPOBA STANDARD; PRT; 608 AA.
AC Q42857;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
DE NCBI_TaxID=4120;
GN WAXY OR SS67.
OS Ipomoea batatas (Sweet potato) (Batate).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX NCBI_TaxID=4120;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Tainong 57; TISSUE=Tuberous root;

```

Wang S.-J., Yeh K.W., Tsai C.Y.;  
Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
-!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.  
-!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =  
UDP + ((1,4)-alpha-D-glucosyl)(N+1).  
-!- PATHWAY: Starch biosynthesis.  
-!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).  
-!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCAGEN SYNTHASE  
FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U41126; AAA86423.1; -  
InterPro; IPR001296; Glycos\_transf\_1.  
Pfam; PF00534; Glycos\_transf\_1; 1.  
Glycogen biosynthesis; Transferase; glycosyltransferase;  
Transit peptide; Chloroplast; starch biosynthesis.  
TRANSIT 1 76 CHLOROPLAST (BY SIMILARITY).  
CHAIN 77 608 GRANULE-BOUND GLYCAGEN [STARCH] SYNTHASE.  
FT BINDING 96 96 UDP-GLUCOSE (BY SIMILARITY).  
SQ SEQUENCE 608 AA; 67000 MW; CCA8FA50A2F69CB0 CRC64;

Query Match 41.9%; Score 1412.5; DB 1; Length 608;  
Best Local Similarity 53.5%; Pred. No. 3.2e-77;  
Matches 292; Conservative 61; Mismatches 150; Indels 41; Gaps 10;

```

QY 2 LDIVYAAEYAPWPKTGGLGVTGGLPTELVKRGHRVMTIAPRYDQYADAWDTSVVVDI- 60
DB 82 ANLVFVGAEGPWSKTGGLGVLGGLPPVLAGNHRVMTSPRYDQYKDAWDTNVLVEVK 141
QY 61 MG-----EKVRFHSIKKGVRWIDHPWFELAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117
DB 142 VGDRIETVRPHCYKRGVDRVDFVDPHLEFLEKRWGKTGSKLYGPKTGIDYRDNQLRFSLLC 201
QY 118 KAATEARVLPF-----GP-GEDCVFVANDHNSALVPVLLKDEYQPKGFTAKSVLA 169
DB 202 QAALAPRVNLNLSNKKYFSGPYGEDVIFVANDHNSALIPCVLKSMYKSRGLYKNAKVAFC 261
QY 170 IHNIAFGRWEEAFKDTKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTGTGY 229
DB 262 IHNIAFGRWEEAFKDTKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTGTGY 304
QY 230 NMLKGGIIAADKLVTSPNYATEIADAAGGVLDTVIRAKGIEGVNGMDIEWNP 289
DB 305 NMLKGGIIAADKLVTSPNYATEIADAAGGVLDTVIRAKGIEGVNGMDIEWNP 289
QY 290 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFTIGRLEQGVDIILAALPKI 349
DB 365 KFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFTIGRLEQGVDIILAALPKI 349
QY 350 LATPKVQVQIALIGTGKAAEKLIVNAIGTKYGRAGVYVFSAPLAHMLTAGADFMVLP 409
DB 425 ISM-DVQIILLTGKTKKFEQIQLEFVMTDPDKARGVAKFNVLPAHMLTAGADFMVLP 483
QY 410 EPCGLIQLHAMHYGTVPVASTGGLVDTVKRGVTFGRHMGALNP--DKLDEADADALA 467
DB 484 EPCGLIQLHAMHYGTVPVASTGGLVDTVKRGVTFGRHMGALNP--DKLDEADADALA 467
QY 468 RRASEVYAGRYPEMVANCISQDLSWSKPAQKWEGLL-----BEVYKGGVATAK 519
DB 544 GRALAIYGLTAPTEMKNCMSQELSKWGPKNWETVLLSIVAGSPGVEGE-EIAPLAK 602
QY 520 EIKVVP 525
DB 603 ENVATP 608

```

## RESULT 6

```

UGST_ORYSA
ID UGST_ORYSA STANDARD; PRT; 609 AA.
AC P19395; O43013;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY OR WX.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=92322986; PubMed=1377969;
RA Okagaki R.J.;
RT "Nucleotide sequence of a long cDNA from the rice waxy gene.";
RL Plant Mol. Biol. 19:513-516(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Japonica Taichung 65; TISSUE=Seedling;
RA Hirano H.Y., Sano Y.;
RT "Molecular characterization of the waxy locus of rice (Oryza
RT sativa)";
RL Plant Cell Physiol. 32:989-997(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Japonica, and cv. Hanfeng;
RX MEDLINE=91016948; PubMed=2216792;
RA Wang Z.Y., Wu Z.L., Xing Y.Y., Zheng F.G., Guo X.L., Zhang W.G.,
RA Hong M.M.;
RT "Nucleotide sequence of rice waxy gene.";
RL Nucleic Acids Res. 18:5898-5898(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica;
RA Wang X.Q., Wang Z.Y., Hong M.M.;
RT Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 153-343 FROM N.A.
RX MEDLINE=91200672; PubMed=2016064;
RA Shimada H., Tada Y.;
RT "Rapid isolation of a rice waxy sequence: a simple PCR method for the
RT analysis of recombinant plasmids from intact Escherichia coli
RT cells.";
RL Gene 98:243-248(1991).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) -
CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
CC -!- PATHWAY: Starch biosynthesis.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62134; CAA44065.1; -
CC EMBL; X58228; CAA41186.1; -
CC EMBL; X53694; CAA37732.1; -
CC EMBL; X65183; CAA46294.1; -
CC EMBL; M55039; AAA33918.1; -
CC PIR; JQ0703; JQ0703.
CC PIR; S22519; S22519.

```

```

DR PIR; S30485; S30485.
DR PIR; JQ2224; JQ2224.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Transit peptide; Chloroplast; Starch biosynthesis.
FT TRANSIT 1 77 CHLOROPLAST.
FT CHAIN 78 609 GRANULE-BOUND GLYCOGEN [STARCH] SYNTHASE.
FT BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).
FT CONFLICT 247 247 N -> T (IN REF. 5).
FT CONFLICT 250 250 P -> T (IN REF. 5).
FT CONFLICT 415 415 P -> S (IN REF. 4).
SQ SEQUENCE 609 AA; 66476 MW; C225DBF6F12072C5 CRC64;

Query Match 41.5%; Score 1397.5; DB 1; Length 609;
Best Local Similarity 51.4%; Pred. No. 2.5e-76;
Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;

QY 2 LDIWVAAEVPWPKTGGGLGDTGGTGLPIELVKGRHVRMTIAPRYDOYADAWDTSVVVDI- 60
Db 83 MNVVFVGAEMAPWPKTGGGLGDTGGTGLPPAMAANGHRVMTISPRDQTKDADWISVVAEIK 142
QY 61 ---MGEKVRVYFHSITKGVHRVWIDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117
Db 143 VADRYEVRFFHCYKRGVDRVFDHPSPLEKVGKTGKIYGPDTGVYDKDNQMRFSLLC 202
QY 118 KAAIEARVL-----PFGP---GEDCVFVANDHNSALVPVLLKDEYQPKGQFTKAKSVLA 169
Db 203 QAALEAPRIILNINNPNYKGTGYGVVFCNDWHTGPLASYLNKNNYQPNGIYRNKAVAF 262
QY 170 IHNIAFGRMWEAEAFKDKLPQAAFDKLAISDGVAKVYVTEATPMEDEKPLTGKTYKKI 229
Db 263 IHNISYQGFADFDYPELNLSERFSRSDFDIDGY-----DTPVEG-----RKI 305
QY 230 NWLKGITIAADKLVTSPNVAETIAADAAGVELDTVIRAKIGIEGIVGMGDIENPNKTD 289
Db 306 NWMKAGILEADRVLTSPVYAEELISGIARCELNDIMRLTGITGIVNGMDVSEWDFPSK 365
QY 290 KFLSAPYDQNSVYAGKAAKEALQALGLPVDPTAPLFAFGRLEEKQGVDDIIAALPKI 349
Db 366 KYITAKYDATTATEAKALKEALQAEAGLPVDRKIPLIAFTGRLEEKQGVDPVMAAIPEL 425
QY 350 LATPKVOJAILGTGKAAYEKLIVNAICTYKGRAGVVKFSAFLAHLTAGADFMIVPSRF 409
Db 426 M-QEDVQIVLLGTGKKKEKLLKSMEEKYKPGKAVVKNFNAFLAHLTAGADFMIVPSRF 484
QY 410 EPCGLIOLHAMHYGTVPVWASTGGLVDTVKEGVTGFHMGALNPD--KLDEADADALAATY 467
Db 485 EPCGLIOLQNMRYGTPCACASTGGLVDTVIEGKTGFHMGRLSVDCXVVEPSDKKVAATL 544
QY 468 RRASEVFAAGGRYPPEWVANCISODLSWSPKPAOKWGLLEEVVYKGG-----VATAK 518
Db 545 KRAIKVVGTPAYEEMVRNMCNODLSWSPKPAOKWGLLEEVVYKGG-----VATAK 602
QY 519 KEETKVP 525
Db 603 KENVAAP 609

RESULT 7
UGST_HORVU
ID UGST_HORVU STANDARD; PRT; 603 AA.
AC P09842;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.

```

OX RN NCBI\_TaxID=4513;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Vogelsanger Gold; TISSUE=Leaf;  
 RX MEDLINE=88303345; PubMed=2970062;  
 RA Rhode W., Becker D., Salamini F.;  
 RT "Structural analysis of the waxy locus from Hordeum vulgare.";  
 RL Nucleic Acids Res. 16:7185-7186(1988).  
 RN [2]  
 RP SEQUENCE OF 76-89.  
 RC STRAIN=cv. H354-295-2-5; TISSUE=Starchy endosperm;  
 RX MEDLINE=94170739; PubMed=8125056;  
 RA Flengsrud R.;  
 RT "Separation of acidic barley endosperm proteins by two-dimensional electrophoresis.";  
 RL Electrophoresis 14:1060-1066(1993).  
 CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.  
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) -  
 CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).  
 CC -!- PATHWAY: Starch biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: X07931; CAA30755.1; -;  
 CC EMBL: X07932; CAA30756.1; -;  
 CC PIR: S01727; YUBHY.  
 CC InterPro: IPR001296; Glycos\_transf\_1.  
 CC Pfam: PF00534; Glycos\_transf\_1; 1.  
 CC Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 CC Transit peptide; Chloroplast; Starch biosynthesis.  
 CC TRANSIT 1 72 CHLOROPLAST (BY SIMILARITY).  
 CC CHAIN 73 603 GRANULE-BOUND GLYCOSYL [STARCH] SYNTHASE.  
 CC BINDING 90 90 UDP-GLUCOSE (BY SIMILARITY).  
 CC SEQUENCE 603 AA; 66211 MW; 0B0B3DE6A8217934 CRC64;  
 CC  
 CC Query Match 41.4%; Score 1395.5; DB 1; Length 603;  
 CC Best Local Similarity 53.3%; Pred. No. 3.3e-76;  
 CC Matches 277; Conservative 71; Mismatches 141; Indels 31; Gaps 7;  
 QY 2 LDIWVAAEVAPWSKGGGLGDLVGGPLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDIM 61  
 Db 76 MNLVFGAEMAPWSKGGGLGDLVGGPLPAPMAANGHRVMTSPRYDQYKDAWDTSVVSEIK 135  
 QY 62 ----GEKVRYFHSIKKGVHRVWDHPWFLAKVWGKTSGLYGRSGADYLDNKRFPALFC 117  
 Db 136 VADEYRVERFFHCYKRGVDRVFDHPWFLKRGKTKYIYGLDAGTDYEDNQRFSLIC 195  
 QY 118 KAAIEAARYL----PF--CP-GEDEVFANDWHSALVPVLKDEYQPGQFTKASVLA 169  
 Db 196 QALAEAPRLNLNNNPFFSGPYGDEVFVFCNDWHTGLLACYLSKNSYQSGIYRTAKVAF 255  
 QY 170 IHNIAFGQRMWEAFKDTKLPOAFAFDKLAFCSDYAKVYTEATPMEDEXPLTGKTKYKI 229  
 Db 256 IHNISVQGRFSDDDFAQLNLPDRFSFFDIDGYK-----PVEG----RKI 298  
 QY 230 NMLKGIITAAKLVTVSPNYATEIAADAAGGVDELDTVIRAKGIEGIVNGMDLEWNPOTD 289  
 Db 299 NNMKAGILQADVLAVTSPYIAELTSGARGCGLDNIMRLTGTITGVNMDVSEWDPKTD 358  
 QY 290 KFLSAPYDONSVYAGKAAKALQALGELGPVDPPTAPLFAFICRLEEQKGVLDILAALPKI 349  
 Db 359 KFLAVNYDITTALEKALNKEALQAEVGLPVDKRPVAFICRLEEQKGPVYIAALPEI 418  
 QY 350 LATPKVQIAILGTGKAAAEKLVNAICTKYKGRAGKGVKFSAPLAHMLTAGADFMVLPSRF 409

Db 419 LKEEDVQIILLGTGKKEKLKSMEEKPGKVRVAVRENAFLAHOMAGADLLAVTSRF 478  
 QY 410 EPCGLLIQIHAMHYGVFPVAVSTGGLVDIVKGVTFHMGALNPD--KIDEADADALAATV 467  
 Db 479 EPCGLLIQIOMGRITGPGPCVASTGGLVDIVKGVTFHMGALNPD--KIDEADADALAATV 538  
 QY 468 FRASEVFAGRGYPMVAVNCISQDLSWSPKPAOKWEGLEEV 507  
 Db 539 KRNVKVGTFPAYQEMVKNCMIQDLSWSPKPAOKWEDVLEL 578  
 RESULT 8  
 ID UGST\_ORYGL STANDARD; PRT; 609 AA.  
 AC 042968;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Granule-bound glycogen [starch] synthase, chloroplast precursor  
 DE (EC 2.4.1.11).  
 GN WAXY.  
 OS Oryza glaberrima (African rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Ehrhartoideae; Oryzaceae; Oryza.  
 CC NCBI\_TaxID=4538;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=cv. GMS1.  
 CC MEDLINE=92134825; PubMed=1685658;  
 CC Umeda M., Ohtsubo H., Ohtsubo E.;  
 CC "Diversification of the rice Waxy gene by insertion of mobile DNA elements into introns.";  
 CC Jpn. J. Genet. 66:569-586(1991).  
 CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.  
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + {(1,4)-alpha-D-glucosyl}(N) -  
 CC UDP + {(1,4)-alpha-D-glucosyl}(N+1).  
 CC -!- PATHWAY: Starch biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: D10472; BAA01272.1; -;  
 CC InterPro: IPR001296; Glycos\_transf\_1.  
 CC Pfam: PF00534; Glycos\_transf\_1; 1.  
 CC Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 CC Transit peptide; Chloroplast; Starch biosynthesis.  
 CC TRANSIT 1 77 CHLOROPLAST (BY SIMILARITY).  
 CC CHAIN 78 609 GRANULE-BOUND GLYCOSYL [STARCH] SYNTHASE.  
 CC BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).  
 CC SEQUENCE 609 AA; 66475 MW; C228BFB9C407FA5 CRC64;  
 CC  
 CC Query Match 41.4%; Score 1394.5; DB 1; Length 609;  
 CC Best Local Similarity 51.4%; Pred. No. 3.8e-76;  
 CC Matches 281; Conservative 75; Mismatches 148; Indels 43; Gaps 9;  
 QY 2 LDIWVAAEVAPWSKGGGLGDLVGGPLPIELVKRGHRVMTIAPRYDOYADAWDTSVVVDI- 60  
 Db 83 MNLVFGAEMAPWSKGGGLGDLVGGPLPAPMAANGHRVMTSPRYDQYKDAWDTSVVSEIK 142  
 QY 61 ----MEKVRYFHSIKKGVHRVWDHPWFLAKVWGKTSGLYGRSGADYLDNKRFPALFC 117  
 Db 143 VADRYRVERFFHCYKRGVDRVFDHPWFLKRGKTKYIYGLDAGTDYEDNQRFSLIC 202





```

RN SEQUENCE FROM N.A.
RC STRAIN=cv. 12311; TISSUE=Seed;
RA Hsing Y.C., Liu C., Yu H., Hsieh J.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) -
CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
CC -|- PATHWAY: Starch biosynthesis.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U23945; AAC49804.1; -.
CC InterPro: IPR001296; Glycosyl_transf_1.
CC Pfam: PF00534; Glycosyl_transf_1; 1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT 1 77 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 78 608 GRANULE-BOUND GLYCOSYL SYNTHASE.
CC BINDING 97 97 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 608 AA; 66074 MW; C31333FA87D2D8A6 CRC64;
CC
CC Query Match 40.2%; Score 1355.5; DB 1; Length 608;
CC Best Local Similarity 50.6%; Pred. No. 8.2e-74;
CC Matches 276; Conservative 76; Mismatches 154; Indels 39; Gaps 11;
CC
CC QY 2 LDIVMVAEAPWSKTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVYDI- 60
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 82 MNVVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVSPRYDQKAWDTSVYSEIK 141
CC QY 61 MG---EKVRYFHSIKGVHRVWIDHPWFLAKVWGKGTGSKLYGPRSGADYLDNHRKRALFC 117
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 142 MGDGYETVRFFHCYKRGVDRVFDHPLFLFRVWGKTEEKIYGPDAITGDKNQLRFSLLC 201
CC QY 118 KAAIEAARVL-----PF--GP-GEDEVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 202 QAALEAPRILSLNNNPFSGPYGVEDVFCNDWHTGFLSCYLSKNSYQSNIGYDKDAFTAF 261
CC QY 170 IHNIAFQGRWEEAFKDTKLPOAFAFDKLAISDGYAKVYTEATPMEDEKPPLTGKTKYKKI 229
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 262 IHNISYQGRFATSDPELNLPERFKSSDFIDGYEK-----PVGE---RKI 304
CC QY 230 NWLKGIIAADKLIVTSPNYATEIAADAAGGVLDIVIRAKGIEGIVNGMDTEENPKTD 289
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 305 NNMKAGILEADRLIVSPVYAEELISGIARGCELDNIMRLTGITGIVNGMDVSEWDPKSD 364
CC QY 290 KFLSAPYDQNSVYAGKAAKALQAEGLPVPDPTAPLFAFISGLEKQKGVDIILALPKI 349
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 365 KYIAVKYDVSTAVEAKALNKEALQAEVGLPVDKRIPLVAFISGLEKQKGPDMVMAAIP-L 423
CC QY 350 LATPKVQIAILGTGKAAEKLVAICTIKYKRAKGVVFKFSAPLAHMLTAGADFMVPSRF 409
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 424 LMEEDTQIVLLGTGKKPERMLMSAEKYPDKVRVAVKFNALAHHIMAGADALLVTSRF 483
CC QY 410 EPCGLIQLHAMHYGVVPVASTGGVLIVTVKGVTFGHHGALNPD--KLDEADADALAAIV 467
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 484 EPCGLIQLQGMRYGTPPCASTGGVLVDITIEGKTGFHMGRLSVDNCVVPADVKKVATTL 543
CC QY 468 RRASEVFAAGRYPENVANCISODLSWSPKPAOKWEGLLEE--VYVGKG-----VATAKE 520
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 544 KEAIVGVGTPTAEEVMKNCMIQDLSWKGPAKNWENVLLSLGVAGGEPGIEGEIAPLAKE 603
CC QY 521 ELKVP 525
CC Db 604 NVAAP 608

```

```

RESULT 11
UGST_MAIZE STANDARD; PRT; 605 AA.
AC PG04713;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Granule-bound glycogen [starch] synthase, chloroplast precursor
DE (EC 2.4.1.11).
GN WAXY.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Kioesgen R.B., Gierl A., Schwarz-Sommer Z., Saedler H.;
RT "Molecular analysis of the waxy locus of Zea mays.";
RL Mol. Gen. Genet. 203:237-244(1986).
CC -|- FUNCTION: REQUIRED FOR THE SYNTHESIS OF AMYLOSE IN ENDOSPERM.
CC -|- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =
CC UDP + ((1,4)-alpha-D-glucosyl)(N+1).
CC -|- PATHWAY: Starch biosynthesis.
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST, GRANULE-BOUND.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOSYL SYNTHASE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X03935; CAA27574.1; -.
CC EMBL: M24258; AAA33520.1; -.
CC PIR: S07314; S07314.
CC MaizeDB; 15806; -.
CC InterPro: IPR001296; Glycosyl_transf_1.
CC Pfam: PF00534; Glycosyl_transf_1; 1.
CC Glycogen biosynthesis; Transferase; Glycosyltransferase;
CC Transit peptide; Chloroplast; Starch biosynthesis.
CC TRANSIT 1 72 CHLOROPLAST.
CC CHAIN 73 605 GRANULE-BOUND GLYCOSYL SYNTHASE.
CC BINDING 91 91 UDP-GLUCOSE (BY SIMILARITY).
CC SEQUENCE 605 AA; 65966 MW; 137F15207DBFC189 CRC64;

```

```

Query Match 40.2%; Score 1353.5; DB 1; Length 605;
Best Local Similarity 50.4%; Pred. No. 1.1e-73;
Matches 275; Conservative 79; Mismatches 153; Indels 39; Gaps 11;
CC
CC QY 2 LDIVMVAEAPWSKTGGGLDVTGGLPIELVKRHRVMTIAPRYDOYADAWDTSVVYDI- 60
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 77 MNVVFVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVWVSPRYDQKAWDTSVYSEIK 136
CC QY 61 MG---EKVRYFHSIKGVHRVWIDHPWFLAKVWGKGTGSKLYGPRSGADYLDNHRKRALFC 117
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 137 MGDGYETVRFFHCYKRGVDRVFDHPLFLFRVWGKTEEKIYGVAGTDYDQNRFLSLLC 196
CC QY 118 KAAIEAARVL-----PF--GP-GEDEVFVANDWHSALVPVLLKDEYQPKGQFTKAKSVLA 169
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 197 QAALEAPRILSLNNNPFSGPYGVEDVFCNDWHTGFLSCYLSKNSYQSNIGYDKDAFTAF 256
CC QY 170 IHNIAFQGRWEEAFKDTKLPOAFAFDKLAISDGYAKVYTEATPMEDEKPPLTGKTKYKKI 229
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 257 IHNISYQGRFATSDPELNLPERFKSSDFIDGYEK-----PVGE---RKI 299
CC QY 230 NWLKGIIAADKLIVTSPNYATEIAADAAGGVLDIVIRAKGIEGIVNGMDTEENPKTD 289
CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 300 NMKAGIIEADRVLTSPYABELISGIARGCELDNIMRLTGITGIVNGMDVSEWDPGRD 359  
 QY 290 KFLSAPYDONSUYAGAAKALQAEGLGVDPDTAPLFAFTGRLEBQKGVDIILALPKI 349  
 Db 360 KIIAVKYDVSTAVEKALKALQAEVGLPVDRNPLVAFITGRLEBQKGVDPVMAAIPOL 419  
 QY 350 L-ATPKVOITAILGTGKAAYEKLVNAIGTKYKGRAGVVKFSAPLAHMTAGADFMVPSR 408  
 Db 420 MEMVEDVOITLITCKKKFERMLMSAEKFKPKVRVAVKVFNAALAHIMAGADVLAVTSR 479  
 QY 409 FEPGGLIOLHAMHYGTVPVASTGGLVDTVKEGVTGFFHMGALNPD--KLDEADADALAT 466  
 Db 480 FEPGGLIOLQGRYGTCPACASTGGLVDTIIEGKTGFHMGRLSVDCNVVPEADVKKVAT 539  
 QY 467 VRRASEVAGGYPWVANCISQDLSWSKPAQKWEGLLEE--VYKGGGV-----ATAKK 519  
 Db 540 LORALKVVGTPAYEEMVRNCMTQDLSWKGPAKNWENVLISLVAGGEPGVEGEIAPLAK 599  
 QY 520 EIKVP 525  
 Db 600 ENVAAP 605

## RESULT 12

UGS2\_SOLU STANDARD; PRT; 641 AA.  
 AC P93568;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Soluble glycogen [starch] synthase, chloroplast precursor  
 DE (EC 2.4.1.11) (SS I).  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RC STRAIN=cv. Desiree; TISSUE=Leaf;  
 RX MEDLINE=97164391; PubMed=9011082;  
 RA Abel G.J.W., Springer F., Willmitzer L., Kossmann J.;  
 RT "Cloning and functional analysis of a cDNA encoding a novel 139 kDa  
 starch synthase from potato (Solanum tuberosum L.).";  
 RL Plant J. 10:981-991(1996).  
 CC -|- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =  
 UDP + ((1,4)-alpha-D-glucosyl)(N+1).  
 CC -|- PATHWAY: Starch biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE (BY  
 SIMILARITY).  
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE  
 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Y10416; CAA71442.1;  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR Pfam; pf00534; Glycos\_transf\_1; 1.  
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 FT TRANSIT ? CHLOROPLAST (POTENTIAL).  
 FT CHAIN ? 641 SOLUBLE GLYCOGEN [STARCH] SYNTHASE.  
 FT BINDING 145 145 UDP-GLUCOSE (BY SIMILARITY).  
 SQ SEQUENCE 641 AA; 70608 MW; 30F8B0546CEFB74C CRC64;

Query Match 26.8%; Score 902.5; DB 1; Length 641;  
 Best Local Similarity 39.8%; Pred. No. 9.9e-47;

Matches 212; Conservative 90; Mismatches 161; Indels 69; Gaps 19;  
 QY 3 DIVMAAEVAPSKTGGLDGVTTGGLPIELVKGRHVMYTIAPRY-----DOYADAWDTS 55  
 Db 132 NIIFVTAEAPYSKGTGLDGVCSLPMALAAHGRVWVSPRYLNGPSPDEKYANAVDL 191  
 QY 56 V--VDIMG--EKVRFHSIKKGVHVWIDHFWFLAKVWGKTGSKLYGSPSGADYLDNHK 111  
 Db 192 VRAIVHCFDAOEAFVFEYRAGVDWVFDHSSY-----RPGTP--YGDYIGA-FGDNOF 244  
 QY 112 REALFCKAIEAARVLPEGP--GDCVFVANDHWSALVPVLIKDEYKQKGTAKSVL 168  
 Db 245 RFTLLSHAACEAPLVLPGLGFTYGEKCLFLANDHAAVPLLLAAKYRPGVYKARSIV 304  
 QY 169 AIHNTIAFOGRMEEAFKDKLPQAFKDLAFSDGYAKVYTEATPMEDEKPPITGTYKK 228  
 Db 305 AIHNTIAHQGVPAVTYNNLGLPPQWYGAV---EWIETWARAHLD-----TGET--- 351  
 QY 229 INWLKGIITAAADKLVTVSPNYATEIAADAAGGVLEDTVI--RAKGTGIVNGMDIEWNP 286  
 Db 352 VNYLKGAIAVADRILTVSOGYSWEIIT-PEGGYLHELLSSROSVLNGTINGIDVNDWP 410  
 QY 287 KTDKFLSAPYDQNSVYAGKAAKEALQAEGLPVDPTAPLFAFTGRLEBQKGVDIILAL 346  
 Db 411 STDEHTASHYSINDL-SGKVQCKTDLQKELGLPIRDCPLIGFIRGLDYKQGVDIILSAI 469  
 QY 347 PKILATPKVOITAILGTGKAAYEKLVNAIGTKYKGRAGVVKFSAPLAHMTAGADFMVLP 406  
 Db 470 PELMQN-DVGQVMLGSGEKQYEDWMRHTENLFKDFRAWVGFNPVSHRTAGCDIILMP 528  
 QY 407 SREPCGLIOLHAMHYGTVPVASTGGLVDTV-----EGV---TGFHMGALNPDKLD 456  
 Db 529 SREPCGLNOLYAMRYGTPIVHSTGGRLDITDFNPYAOEGEGEGTGTWTFSPITSKL- 587  
 QY 457 EADADALAAVRRASEVFAAGRYPE-----MVANCISQDLSWSKPAQKWE 501  
 Db 588 ---LDILKLA-----GTYTEHKSWEGLMRMGGRDYSWENAAIQYE 627

## RESULT 13

UGS2\_ORYSA STANDARD; PRT; 626 AA.  
 AC Q40739;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Soluble glycogen [starch] synthase, chloroplast precursor  
 DE (EC 2.4.1.11) (SSS).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RZ SEQUENCE FROM N.A., AND SEQUENCE OF 114-131.  
 RC STRAIN=cv. Japonica; TISSUE=Seed;  
 RX MEDLINE=94302151; PubMed=7518089;  
 RA Baba T., Nishihara M., Mizuno K., Kawasaki T., Shimada H.,  
 RA Kobayashi E., Ohnishi S., Tanaka K.-I., Arai Y.;  
 RT "Identification, cDNA cloning, and gene expression of soluble starch  
 synthase in rice (Oryza sativa L.) immature seeds.";  
 RL Plant Physiol. 103:565-573(1993).  
 CC -|- CATALYTIC ACTIVITY: UDP-glucose + ((1,4)-alpha-D-glucosyl)(N) =  
 UDP + ((1,4)-alpha-D-glucosyl)(N+1).  
 CC -|- PATHWAY: Starch biosynthesis.  
 CC -|- SUBCELLULAR LOCATION: AMYLOPLAST OR CHLOROPLAST, SOLUBLE.  
 CC -|- TISSUE SPECIFICITY: LEAVES AND IMMATURE SEEDS.  
 CC -|- MISCELLANEOUS: THREE FORMS OF SOLUBLE STARCH SYNTHASE WERE  
 CC PURIFIED: RSS1, RSS2 AND RSS3.  
 CC -|- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration





*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:35:26 ; Search time 32.3424 Seconds  
(without alignments)  
4147.394 Million cell updates/sec

Title: US-09-980-771A-5  
Perfect score: 3370  
Sequence: 1 ALDIYVAAEYAPWKTGGL.....SASKTSAARPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3366	99.9	708	10	064925
2	1463.5	43.4	613	10	Q93YB1
3	1450.5	43.0	604	10	Q9FR03
4	1424.5	42.3	608	10	Q93VD9
5	1423.5	42.2	606	10	Q9X1S6
6	1423.5	42.2	607	10	Q92SQ5
7	1404	41.7	608	10	Q8VYU1
8	1397.5	41.5	609	10	Q94LY7
9	1395.5	41.4	603	10	Q8SA49
10	1390	41.2	610	10	Q9MAO0
11	1389.5	41.2	609	10	Q8S9C4
12	1381	41.0	605	10	Q8S9K3
13	1380	40.9	605	10	Q9FU06
14	1379	40.9	599	10	Q9SQ58
15	1376	40.8	606	10	Q43012
16	1375	40.8	604	10	Q9S7N5

17	1371	40.7	534	10	Q8W2G8	Q8W2g8
18	1369	40.6	604	10	Q9SL56	triticum tu
19	1367.5	40.6	604	10	Q9SL57	triticum tu
20	1367	40.6	605	10	Q9SQ51	Q9sq51 aegilops sp
21	1366	40.5	605	10	Q9SL58	triticum tu
22	1362.5	40.4	604	10	Q9SXX4	Q9sxx4 triticum ae
23	1360.5	40.4	604	10	Q9SL59	triticum tu
24	1354.5	40.2	605	10	Q9SQ52	Q9sq52 triticum mo
25	1353.5	40.2	574	10	Q9SVU0	Q9svu0 triticum ae
26	1263.5	37.5	565	10	Q9XEN9	Q9xen9 triticum ae
27	898	26.6	792	10	Q9MAC8	Q9mac8 arabidopsis
28	895.5	26.6	313	10	Q9LKD3	Q9lkd3 vaquelinia
29	893.5	26.5	313	10	Q9LKE0	Q9lke0 prunus virg
30	893.5	26.5	313	10	Q9FYU8	Q9fyu8 aruncus dio
31	891.5	26.5	647	10	Q9LEB9	Q9leeb9 triticum ae
32	891.5	26.5	647	10	Q9SQH0	Q9sqh0 aegilops ta
33	891.5	26.5	647	10	Q9SQG9	Q9sqg9 triticum ae
34	890.5	26.4	647	10	Q9LEC0	Q9lec0 triticum ae
35	886.5	26.3	313	10	Q9LKE8	Q9lke8 kageneckia
36	886	26.3	576	10	Q64926	Q64926 chlamydomon
37	885.5	26.3	313	10	Q9LKD1	Q9lkd1 exochorda r
38	884.5	26.2	643	10	Q9M5A3	Q9m5a3 hordeum vul
39	879.5	26.1	313	10	Q9LKE5	Q9lke5 amelanchier
40	879.5	26.1	313	10	Q9LKD6	Q9lkd6 vaquelinia
41	878.5	26.1	313	10	Q9LKE7	Q9lke7 kageneckia
42	878.5	26.1	313	10	Q9LKE1	Q9lke1 prinsepia s
43	877.5	26.0	313	10	Q9LKE6	Q9lke6 malus sarge
44	877.5	26.0	313	10	Q9LKD8	Q9lkd8 rosa multif
45	876.5	26.0	313	10	Q9LKE5	Q9lke5 oemleria ce

#### ALIGNMENTS

RESULT 1  
ID 064925 PRELIMINARY; PRT; 708 AA.  
AC 064925;  
DT 01-AUG-1998 (TREMREL.07, Created)  
DT 01-DEC-2001 (TREMREL.19, Last sequence update)  
DT 01-MAR-2002 (TREMREL.20, Last annotation update)  
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).  
GN STA2.  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=137C;  
RA D'Hulst C., Wattlebled F., Ral J.-P., Abel G.J., Kossmann J.,  
RA Ball S.G.;  
RT "Cloning of a cDNA encoding for the GBSI in the green alga  
RT Chlamydomonas reinhardtii."  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Wattlebled F., Ball S.G., D'Hulst C.;  
RT "Granule-bound starch synthase I: A major enzyme involved in the  
RT biogenesis of B-crystallites in starch granules."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF026420; AAC19969.3; -;  
DR EMBL; AF433156; AAL28128.1; -;  
DR InterPro; IPR001296; Glycos.transf.1.  
DR InterPro; IPR002114; HPT\_Serp\_site.  
DR Pfam; PF00534; Glycos.transf.1.  
DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN\_1.  
KW Glycosyltransferase; Transferrase; TRANSIT.  
FT CHAIN 1 57 POTENTIAL.  
SQ SEQUENCE 708 AA; 74623 MW; 7D2A5A07D8606469 CRC64;

Query Match 99.9%; Score 3366; DB 10; Length 708;

Best Local Similarity 99.8%; Pred. No. 4.5e-190;  
Matches 650; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALDIYVAAEVAPWSKTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 60  
DB 58 ALDIYVAAEVAPWSKTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 117  
QY 61 MGKRVYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNKKRALCKAA 120  
DB 118 MGKRVYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNKKRALCKAA 177  
QY 121 IEAARVLPFGEDCVFVANDWHSALVPVLLKDEYQPKQFTAKSVLAHNTAFQGRMW 180  
DB 178 IEAARVLPFGEDCVFVANDWHSALVPVLLKDEYQPKQFTAKSVLAHNTAFQGRMW 237  
QY 181 EFAKDTKLPOAFAFDKLAESDGYAKVYTEATPMEDEKPLTKGTYKKNLWKGGLIAD 240  
DB 238 EFAKDTKLPOAFAFDKLAESDGYAKVYTEATPMEDEKPLTKGTYKKNLWKGGLIAD 297  
QY 241 KLVTSPNPTATEIAAAGGVELDTVIRAKGIEGIVNGMDIEWNPKTDFLSAPYDQNS 300  
DB 298 KLVTSPNPTATEIAAAGGVELDTVIRAKGIEGIVNGMDIEWNPKTDFLSAPYDQNS 357  
QY 301 VYAGKAAAEALQAEGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 360  
DB 358 VYAGKAAAEALQAEGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 417  
QY 361 GTGKAAEYKLVNAIGTKYGRAGVVKFSAPLAHMLTAGADPMLVPSRPEPCGLIOLHAM 420  
DB 418 GTGKAAEYKLVNAIGTKYGRAGVVKFSAPLAHMLTAGADPMLVPSRPEPCGLIOLHAM 477  
QY 421 HYGTVPVASTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 480  
DB 478 HYGTVPVASTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 537  
QY 481 EMVANCISODLSWSKPAQWEGLEEVYKGGVATAKKEEIKVPVAKETPGDLPVASYA 540  
DB 538 EMVANCISODLSWSKPAQWEGLEEVYKGGVATAKKEEIKVPVAKETPGDLPVASYA 597  
QY 541 PNTLKPVSASVEGNGAAAPKVGTPAPAMGAWRATTSPGSPAAATPKVTYKPAALPATAK 600  
DB 598 PNTLKPVSASVEGNGAAAPKVGTPAPAMGAWRATTSPGSPAAATPKVTYKPAALPATAK 657  
QY 601 PKTAGLKLAGEASTTSTSENGAASNGNGASASKTSAAKPLVSAATRKA 651  
DB 658 PKTAGLKLAGEASTTSTSENGAASNGNGASASKTSAAKPLVSAATRKA 708

RESULT 2  
Q93YB1 PRELIMINARY; PRT; 613 AA.

ID Q93YB1  
AC Q93YB1  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Granule-bound starch synthase (EC 2.4.1.11).  
GN GBSSI.  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LEAF;  
RA Edwards A., Vincken J.P., Visser R., Zeeman S., Smith A.M., Martin C.;  
RT "Discrete forms of amylose are synthesised by isoforms of GBSSI in  
pea.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ345045; CAC69955.1; -;  
DR InterPro; IPR001296; Glycos\_transf\_1.  
DR InterPro; IPR001917; Ntrtransf\_2.  
DR Pfam; PF00534; Glycos\_transf\_1; -.

DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; UNKNOWN\_1.  
KW Glycosyltransferase; Transferase; Transf. No. 4.1e-78;  
FT TRANSIT 1 85  
FT CHAIN 86 613 GRANULE BOUND STARCH SYNTHASE.  
SQ SEQUENCE 613 AA; 67626 MW; 5A653461128F97EB CRC64;

Query Match 43.4%; Score 1463.5; DB 10; Length 613;  
Best Local Similarity 53.0%; Pred. No. 4.1e-78;  
Matches 291; Conservative 74; Mismatches 137; Indels 47; Gaps 11;

QY 2 LDIVMVAEAVAPWSKTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI- 60  
DB 87 MNIFVGTVEAPWSKTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 146  
QY 61 --MGKRVYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNKKRALFC 117  
DB 147 VGRTEKRVFFHCFKRGVDRVFDHPFIFLEKVGKGTGKLYGAAGDDYQDNLRSIFC 206  
QY 118 KAAIEAARVLPF-----GP-GEDEVFVANDWHSALVPVLLKDEYQPKQFTAKSVLA 169  
DB 207 QAAIEAARVLPF-----GP-GEDEVFVANDWHSALVPVLLKDEYQPKQFTAKSVLA 266  
QY 170 IHNIAFGQRMWEAFKD---TKLPQAAFDKLAESDGYAKVYTEATPMEDEKPLTKGTY 226  
DB 267 IHNIAFGQRMWEAFKD---TKLPQAAFDKLAESDGYAKVYTEATPMEDEKPLTKGTY 306  
QY 227 KKNLWKGGLIADKLAESDGYAKVYTEATPMEDEKPLTKGTYKKNLWKGGLIAD 286  
DB 307 KKNLWKGGLIADKLAESDGYAKVYTEATPMEDEKPLTKGTYKKNLWKGGLIAD 366  
QY 287 KTKFSLAPYDQNSVYAGKAAAEALQAEGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 346  
DB 367 STDKYISIKYDASTVLEKALKEELQAEVCLPVDKNVPLIAFIFGLEEQKSDILVFAI 426  
QY 347 PKILATPKVOIATILGTCVAAEYKLVNAIGTKYGRAGVVKFSAPLAHMLTAGADPMLV 406  
DB 427 PQFI-KENVQIVAGLTKGKEMEKQLOLEISYDPDKARGVAKFNVPLAHMIAAGDFILIP 485  
QY 407 SRPEPCGLIOLHAMHYGTVPVASTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 464  
DB 486 SRPEPCGLIOLHAMHYGTVPVASTGLGSDVDTGGLPIELVYKGRHVMTIAPRYDOYADAWDTSVVVDI 545  
QY 465 ATYRRASEVTFAGRYDEMVAANCISODLSWSKPAQWEGLEEVYKGGVATAKKEEIKVP 516  
DB 546 KTYTKALGVYGTSAFAEMIKNCAQELSWKGPAPKKEEVLNLLGVDPSEPDGQ-ETAP 604  
QY 517 AKKEIKVP 525  
DB 605 QAKENATP 613

RESULT 3  
Q9FR03 PRELIMINARY; PRT; 604 AA.

ID Q9FR03  
AC Q9FR03  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Granule-bound starch synthase.  
GN GBSSI.  
OS Perilla frutescens.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.  
OX NCBI\_TaxID=48386;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. OKDONG; TISSUE=DEVELOPING SEEDS;  
RA Hwang S.-K., Hwang Y.-S.;  
RT "Isolation and characterization of a cDNA encoding granule-bound  
starch synthase from Perilla frutescens.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF210699; AAG43519.1; -.



```

DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
SQ SEQUENCE 604 AA; 66214 MW; 435CE9601C85A5EF CRC64;

Query Match      43.0%; Score 1450.5; DB 10; Length 604;
Best Local Similarity 54.3%; Pred. No. 2.3e-77;
Matches 297; Conservative 62; Mismatches 145; Indels 43; Gaps 10;

QY 2 LDIVMTAAEAVAPSKTGGGLGDTVGGPLPIELVKGRHVRMTIAPRYDQYADAWDTSVVVDI- 60
Db 78 MTLIFVSAEGPWSKTGGGLGDTVGGPLPALAANRHRVMTSPRYDQYADAWDTSVVVDI- 137
QY 61 MGEKVR---YFHSIKKGVHRVWDHPWFLAKVWGKTGSKLYGPRSGADYLDNKRKRALFC 117
Db 138 VGDKVTVEGFFHCYKRGVDRVFDHPLFLEKVGKTKSKVYGSAGVDYEDNQLRFLSLS 197
QY 118 KAIEAARVLFP-----GP-GEDCVFVANDWHSALVPLVLLKDEYOPKGOFTKAKSVLA 169
Db 198 LAALAPRVNLTSNKFSGPYGEDVVFVANDWHTAIVLPCYLKTIYQPKGIYNNAKVLC 257
QY 170 IHNIAFGRMWEAEAFKDTKLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPLTKYTKKI 229
Db 258 IHNIAFGRAFSDFLKLNLPDQLKSSDFMDGYEK-----PVKG---RKI 300
QY 230 NWLKGITIAADKLVTVSPNATETIAADAAGGVLDTVIRAKGTGIVNGMDIEWNPKTID 289
Db 301 NNMKAGIESDRVLTVSPYANELVSGPKGVLDNLRKCTVIGVINGMDTQEWNPATD 360
QY 290 KFLSAPYDQNSVYAGAAAEALQAEGLPDPVTPALFAPIGLEBQKGVDTIILALPKI 349
Db 361 KYIDNHYDITVMDGKPLLEALQAEVGLPDRNVPVGVLFGRLEBQKGVDTIILALPKI 420
QY 350 LATPKVQIALLGTGKAAYEKLVNAIGTKYGRAGKGVVVFSAPLAHMLTAGADFMVPSRF 409
Db 421 IEM-DVQVILGTGKKEFEQIEQLEVMYPDKARGVAKFNVPLAHMTAGADFMVPSRF 479
QY 410 EPCGLIQLHAMHYGTVPVASTGGIVDTVKEGTVGFHMGALNP--DKLDEADADALAATV 467
Db 480 EPCGLIQLHAMRYGTIPICASTGGIVDTVKEGTVGFHMGAFNVPECDAPADVLKIVTIV 539
QY 468 RRASEVFAGRPYEMVANCISDLSWSKPAQKWEGLLEEVYVYKGG-----VATAK 518
Db 540 GRALEYIGTPAPFREMNNCSLDLSWKGPAKNWETVL--LSLGVAGSEPGVEGDEIAPLA 597
QY 519 KEIKVP 525
Db 598 KENVATP 604

RESULT 4
Q93VD9 PRELIMINARY; PRT; 608 AA.
AC Q93VD9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Granule-bound starch synthase I (EC 2.4.1.21).
GN GBSI.
OS Ipomoea batatas (Sweet potato) (Batata).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Convolvulaceae; Ipomoea.
OX NCBI_TaxID=4120;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. KOKEI 14;
RA Kimura T., Saito A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. KOKEI 14;
RA Kimura T., Ideta O., Saito A.;
RT *Identification of the gene encoding granule-bound starch synthase I

```

```

RT in sweet potato (Ipomoea batatas (L.) Lam.).";
RL Plant Biotechnol. 17:247-252(2000).
DR EMBL; AB071976; BAB68525.1; -
DR EMBL; AB071604; BAB68126.1; -
DR InterPro: IPR001296; Glycos_transf_1.
DR Pfam: PF00534; Glycos_transf_1; 1.
KW Glycosyltransferase; transferase.
SQ SEQUENCE 608 AA; 66689 MW; C93CDA49A1F50C03 CRC64;

Query Match      42.3%; Score 1424.5; DB 10; Length 608;
Best Local Similarity 53.6%; Pred. No. 7.9e-76;
Matches 293; Conservative 61; Mismatches 150; Indels 43; Gaps 10;

QY 2 LDIVMTAAEAVAPSKTGGGLGDTVGGPLPIELVKGRHVRMTIAPRYDQYADAWDTSVVVDI- 60
Db 82 MNLVFGCEVGPWCKTGGGLGDTVGGPLPALAARHVRVTCPRYDQYADAWDTSVVVDI- 141
QY 61 MG---EKVYFHSIKKGVHRVWDHPWFLAKVWGKTGSKLYGPRSGADYLDNKRKRALFC 117
Db 142 VGDRIEVPREFHSYKRGVDRVFDHPLFLEKVGKTKGSMYLGPKAGKDYDNQLRFLSLLC 201
QY 118 KAIEAARVLFP-----GP-GEDCVFVANDWHSALVPLVLLKDEYOPKGOFTKAKSVLA 169
Db 202 QAALAPRVNLNLSNNTFSGPYGEDVVFVANDWHTALLPCYLKTIYOSRGIYNNAKVAF 261
QY 170 IHNIAFGRMWEAEAFKDTKLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPLTKYTKKI 229
Db 262 IHNIAFGRAFSDFLKLNLPDQKSGDFIDGYK-----PVKG---RKI 304
QY 230 NWLKGITIAADKLVTVSPNATETIAADAAGGVLDTVIRAKGTGIVNGMDIEWNPKTID 289
Db 305 NNMKAGITREADRYTVSPYAKELVSCVSGVLDNHRDCGTGICNGMDTQEWNPATD 364
QY 290 KFLSAPYDQNSVYAGAAAEALQAEGLPDPVTPALFAPIGLEBQKGVDTIILALPKI 349
Db 365 KYLAVKXDIITVMAQKPLLEALQAEVGLPDRNVPVGVLFGRLEBQKGVDTIILALPKI 424
QY 350 LATPKVQIALLGTGKAAYEKLVNAIGTKYGRAGKGVVVFSAPLAHMLTAGADFMVPSRF 409
Db 425 ISM-DVQIILLGTGKKEFEQIEQLEVMYPDKARGVAKFNVPLAHMTAGADFMVPSRF 483
QY 410 EPCGLIQLHAMHYGTVPVASTGGIVDTVKEGTVGFHMGALNP--KLDADADALAATV 467
Db 484 EPCGLIQLHAMRYGTIPICASTGGIVDTVKEGTVGFHMGAFNVPECDVPELVKIVTV 543
QY 468 RRASEVFAGRPYEMVANCISDLSWSKPAQKWEGLLEEVYVYKGG-----VATAK 518
Db 544 GRALAMYGLAFTTEMKNCMSQELSNKGPAPKNWETVL--LSLGVAGSEPGVEGDEIAPLA 601
QY 519 KEIKVP 525
Db 602 KENVATP 608

RESULT 5
Q9XIS6 PRELIMINARY; PRT; 606 AA.
AC Q9XIS6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Granule-bound starch synthase I precursor (EC 2.4.1.21).
GN GBSI.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N. Nozaki K., Ito H., Matsui H., Honma M.;
RT *Phaseolus vulgaris L. mRNA for Granule-Bound Starch Synthase I
(GBSI).";

```



Query Match	41.7%	DB 1404;	DB 10;	Length 608;
Best Local Similarity	51.5%;	Pred. No. 1.3e-74;		
Matches	283;	Conservative 72;	Mismatches 147;	Indels 48; Gaps 10;
QY	2	LDIVWAAEYAPWSKTGGGLDVTGGGLPIELVYKGRHVRWTTAPRYDQADAMWTSVVDI-	60	
DB	81	MTIIFIATECHPWCKTGGGLDVLGGLPPALAAAGHRVMTIVPRDYDQADAMWTVLVEVN	140	
QY	61	---MGEKRVYFHSIKGVRWIDHPWFLAKRWCKTGSKLYGPRSGADYLDNHRKRALFC	117	
DB	141	IGDRTEVTRFHCYKRGVDRVFDVHPMLEKLVCKTGKLYGPTGGDYDRDNQURFCLLC	200	
QY	118	KAATEARVLPP-----GP-GEDCVFVANDWHSALVPVLKDEYQPKQFTKAKSVLA	169	
DB	201	LAALEAPRVLLNNSEYFSGPYGENYVAVANDWHTGVLPCYLKSIYQAKGYVNAKVAFC	260	
QY	170	IHNIAFOGRWMEAFKDTKLPOAAFDKLAESGAKVYVTEATPMEDEKPLTGKTYKKI	229	
DB	261	IHNIAVOGRAREDFELLNLPDSFLPSFDIDGHFK-----PVG---RKI	303	
QY	230	NWLKGGTIAADKLVTVSPNYATEIAADAAGVELLDTVIRAKGIP-GIVNGMDIEBWNPKT	288	
DB	304	NWMAKAGITECDLVMTVSPHYVKELASGPDGKVELDGLIRTKLETGIVNGMDVYENPAT	363	
QY	289	DKFLSAPYDQNSYVAGAAAKAEOALGELPVPDTPAFIPAFIRGLEQKGVDIILALPK	348	
DB	364	DQIISVYIDATVTTBARALKEMLOAEVGLVDPSSIPLIYFVGRLEPKQKSDILIAAIP	423	
QY	349	ILATPKVQIAILGTGKAAEYKLVNAICTKYKGRAGVVKTSAPLAHMLTAGADMLVPSR	408	
DB	424	FV-EGNVQIIVLGTGKKKEEELILLEVKYPNTARGLAKNVPLAHMFAADFIIVPSR	482	
QY	409	FEPCLGLIQLHAMHYGVVPVASTGGVLVDIVKEGVTGFHMGALNP--DKLDEADADALAAT	466	
DB	483	FEPCLGLIQLQGMRYGVVPICSSTGGVLVDIVKEGVTGFHMGLENVECTVPDVTAVAST	542	
QY	467	VRRASEVFGGRYPMPVANCISODLSKSPAKWEGGLEEVYVYKGG-----GVA	515	
DB	543	VKRALQKYNTPAQEVMQVQNCMAQDLSWKGPAAKW---EEVLGLGVGEGSQPIEGSEVA	598	
QY	516	TAKKEEIKVP 525		
DB	599	PLAKENVATP 608		
RESULT 8				
Q94LY7				
ID	Q94LY7	PRELIMINARY;	PRT;	609 AA.
AC	Q94LY7;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Granule-bound starch synthase (EC 2.4.1.21).			
GN	WAXY.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxId=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. REXMONT.			
RA	Larkin P.D., McClung A.M., Ayres N.M., Park W.D.;			
RT	"The Wx locus (Granule Bound Starch Synthase) is strongly associated			
RL	with pasting curve characteristics in rice (Oryza sativa L.).";			
RL	Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF141954; AAF72561.1;			
DR	InterPro; IPR001296; Glycos.transf_1.			
DR	InterPro; IPR000794; ketoacyl-synt.			
DR	Pfam; PF00534; Glycos.transf_1; 1.			
DR	PROSITE; PS00606; B_KETOTACIL_SYNTHASE; UNKNOWN_1.			
SQ	SEQUENCE 609 AA; 66466 MW; 6D2615EB203CB0D8 CRC64;			

Query Match	41.5%;	Score 1397.5;	DB 10;	Length 609;
Best Local Similarity	51.4%;	Pred. No. 3.1e-74;		
Matches 281;	Conservative 76;	Mismatches 147;	Indels 43;	Gaps 9;
QY	2	LDIYVAAEVPWSTGSLGSDVTGGLPTIELVKRHRVTIAPRYDQYDADWDTSVVDI-	60	
Db	83	MNVFVGAEMAPWSTGSLGSDVGLGPPAMAANGHRVMIIPRYDQYDADWDTSVVAETK	142	
QY	61	--MGEKRYRFSIKKGVHRVWDHPWLAKVWGTSKLYGPRSGADYDLNKHKFAALFC	117	
Db	143	VADRYERVFFHCYKRGYDRVFIIDHPSLEKVGWKTGEKIIGPDGVDYKDNQMFSLIC	202	
QY	118	KAATIAARVL-----PFGP---GEDCVFANDWHSALVPVLLKDEYQKGFOTKAKSVLA	169	
Db	203	QAALFAPRILNNNPYFKGTGVEDVVFVPCNDWHITGGLASLYLKNNYQNGIYRNKAVFC	262	
QY	170	IHTIAFOGRWEEAFKDTKLPOAAFDKLAFLSDGYAKVYTEATPMEDEKPLTGKTYKKI	229	
Db	263	IHTISYOGRFADYPELNUSEFRSSPFDITGTY-----DTPVEG-----RKI	305	
QY	230	NWLKGGIILADKLVTSFNVYATEIAADAAGGVGLDVTIRAKGIEGIVNGMDEEWNPKTD	289	
Db	306	NWMAKILADRLVTSFNYAEELISGTARCELDNMLRTGITGIVNGMDVSEWDSPSKD	365	
QY	290	KFLSAPYDQNSVYAGKAAKALQAEALGLPVDTPAPLFAFAGRLEEKQGVDIILALPKI	349	
Db	366	KYITAKYDATTAEIKALNKALKEALQAEAGLPVDKRIPLIPIAFIGRLEEKQGSVMMAAIPEL	425	
QY	350	LATPKVQIILGTGKAAEKLNVNAIGTKYKGRAGVVKFSAPLAHMLTAGDFMLVPSRF	409	
Db	426	M-QEDVQIVLLGTGKKFKKLLKMEKYPKGVKRVKFNAPLAHLMAGADVLAVPSRF	484	
QY	410	EPCGLIQLHAMHYGTPVPVASTGLVDTKRGVGTGFHMGALNPD--KIDEADADALAATV	467	
Db	485	EPCGLIQLQGMRYGTPPCACASTGLVDTVTEGKTGFHMGRLSDCKVVPESDVKKVAATL	544	
QY	468	RRASEVFAGGRYPENWANCISQDLSWSKPAQKWGLLEEVYVYKGG-----VATAK	518	
Db	545	KAIKVYGTTPAYEEMVRNQMNDLSWKGPAKNWENVL--LGLGVAGSAPGIEGDEIAPLA	602	
QY	519	KEETKVP 525		
Db	603	KENVAAP 609		
RESULT 9				
Q8SA49	ID	Q8SA49	PRELIMINARY;	PRT; 603 AA.
AC	Q8SA49;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Granule-bound starch synthase.			
GN	259116.5			
OS	Hordeum vulgare (Barley).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;			
OC	Triticeae; Hordeum.			
CX	NCBI_TaxID=4513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. MEREX;			
RA	Ma J., SanMiguel P.J., Dubcovsky J., Shiloff B.A., Rostoks N.,			
RA	Jiang Z., Bussio C.S., Kleinhofs A., Devos K.M., Ramakrishna W.,			
RA	Bennetzen J.L.;			
RT	"Comparative sequence analysis of wx1 homologous regions in barley,			
RT	maize, pearl millet, rice, sorghum and diploid wheat."			
RT	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF474373; AAL77109.1;			
SQ	SEQUENCE 603 AA; 66280 MW; 467A3DB6A82125CB CRC64;			
Query Match	41.4%;	Score 1395.5;	DB 10;	Length 603;
Best Local Similarity	53.3%;	Pred. No. 4e-74;		



```

Db 143 VADRYRVRFFHCYKHGVDRVDFIDHPSELEKVGKTEGIYDPTDGVGHKQONQRFSLIC 202
QY 118 KAAEAAARVL-----PFGP-----GECVFAVNDHWSALVPVLLKDEYQPKGQFTAKSVLA 169
Db 203 QAALEAPRILNINNPFYKGTGYGVDFVFCVNDWHTGLASVLLKNNYQPNGIYRNAKVAFC 262
QY 170 IHNIAFGRMWEEAFKDKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTKTKYKKI 229
Db 263 IHNISYQGRFAFEDYPELUSERFRSSDFIDGY-----DTPVEG-----RKI 305
QY 230 NWLKGIIAADKLVTSVSPNYATEIAAAGGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
Db 306 NWMKAGILEADRLVTSVSPYAEELISGTARCELDNIMRLTGITGIVNGMDVSEWDPSKD 365
QY 290 KFLSAPYDONSIVYAKAAKAAKALQAEGLPVDPTAPLFAFTGRLEEQKGVDIILAAALPKI 349
Db 366 KYITAKYDAITTAIEAKALNKEALQAEGLPVDKIPVDRKIPVDRKIPVDRKIPVDRKIPVDR 425
QY 350 LATPKVQITAILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAFLAHLMTAGADFMLVPSRF 409
Db 426 M-QEDVOIVLLGTGKKFERLLKSVYEEKFPKVRVAVRFPNAPLAHLMAGADVLAVTSRF 484
QY 410 EPCGLIOLHAMHYGTVPVASTGGGLVDTVKEGVTGFHMGALNPD--KLDEADADALAATV 467
Db 485 EPCGLIOLQGMRYGTGTPCACASTGGGLVDTVIEGKTGFHMGRLSVDCVKVPSDVKKVAATL 544
QY 468 RRASEVPAGGRYPPEWVANCISODLSWSKPAQKWEGLLEEVYVYKGG-----VATAK 518
Db 545 KRAIKVGTGTAYHEMVKMNCMIODLSWGPKNWENVL--LGLGVAGSAPGTEGDEIAPLA 602
QY 519 KEIKVP 525
Db 603 KENVAAP 609

```

## RESULT 12

```

Q9SXK3 ID Q9SXK3 PRELIMINARY; PRT; 605 AA.
AC Q9SXK3;
DT 01-MAY-2000 (TremBrel. 13, Created)
DT 01-MAY-2000 (TremBrel. 13, Last sequence update)
DE 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE Starch synthase (GBSSI) (EC 2.4.1.21).
GN WAXY.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99321800; PubMed-10393240;
RA Mural J., Taira T., Ohta D.;
RT "Isolation and characterization of the three waxy genes encoding the
RL Gene 234:71-79(1999).
DR EMBL; AB019623; BAA77351.1; -
DR InterPro; IPR001296; Glycosyltransf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycosyltransf_1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 605 AA; 66345 MW; 60A816276F78722B CRC64;

```

```

Query Match 41.0%; Score 1381; DB 10; Length 605;
Best Local Similarity 51.5%; Pred. No. 2.9e-73;
Matches 281; Conservative 72; Mismatches 153; Indels 40; Gaps 9;
QY 2 LDIWVAAEVAEPWMSKYGGLGDTVGGPLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDIM 61
Db 78 MNLVFGAEMAPWSKYGGLGDLVGLGPPAANGHRVMTISPRYDQYKDAWDTSVVSEIK 137

```

```

QY 62 ----GEKRYFHSIKKGVHRVWIDHPFELAKWGKTSKLYGPRSGADYLDNKHKRALFC 117
Db 138 VADEYRVYFHCYKRGVDRVDFIDHPSELEKVGKTEGIYDPTDGVGHKQONQRFSLIC 197
QY 118 KAAEAAARVL-----PFGP-----GECVFAVNDHWSALVPVLLKDEYQPKGQFTAKSVLA 169
Db 198 QAALEAPRILNINNPFYKGTGYGVDFVFCVNDWHTGLASVLLKNNYQPNGIYRNAKVAFC 257
QY 170 IHNIAFGRMWEEAFKDKLPQAAFDKLAFLSDGYAKVYVTEATPMEDEKPLTKTKYKKI 229
Db 258 IHNISYQGRFAFEDYPELUSERFRSSDFIDGY-----DTPVEG-----RKI 300
QY 230 NWLKGIIAADKLVTSVSPNYATEIAAAGGVVELDTVIRAKGIEGIVNGMDIEEWNPKTD 289
Db 301 NWMKAGILEADRLVTSVSPYAEELISGTARCELDNIMRLTGITGIVNGMDVSEWDPAKD 360
QY 290 KFLSAPYDONSIVYAKAAKAAKALQAEGLPVDPTAPLFAFTGRLEEQKGVDIILAAALPKI 349
Db 361 KYITAKYDAITTAIEAKALNKEALQAEGLPVDKIPVDRKIPVDRKIPVDRKIPVDRKIPVDR 420
QY 350 LATPKVQITAILGTGKAAVEKLVNAIGTKYKGRAGVVKFSAFLAHLMTAGADFMLVPSRF 409
Db 421 LKEEDVOIVLLGTGKKFERLLKSVYEEKFPKVRVAVRFPNAPLAHLMAGADVLAVTSRF 480
QY 410 EPCGLIOLHAMHYGTVPVASTGGGLVDTVKEGVTGFHMGALNPD--KLDEADADALAATV 467
Db 481 EPCGLIOLQGMRYGTGTPCACASTGGGLVDTVIEGKTGFHMGRLSVDCVKVPSDVKKVAATL 540
QY 468 RRASEVPAGGRYPPEWVANCISODLSWSKPAQKWEGLLEEV-----VYKGGVATAKK 519
Db 541 KRAIKVGTGTAYHEMVKMNCMIODLSWGPKNWENVL--LGLGVAGSAPGTEGDEIAPLA 599
QY 520 BEIKVP 525
Db 600 ENVAAP 605

```

## RESULT 13

```

Q9FU06 ID Q9FU06 PRELIMINARY; PRT; 605 AA.
AC Q9FU06;
DT 01-MAR-2001 (TremBrel. 16, Created)
DT 01-MAR-2001 (TremBrel. 16, Last sequence update)
DE 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE Granule bound starch synthase I (EC 2.4.1.21).
GN GBSSI.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CV. CHEYENNE; TISSUE-ENDOSPERM;
RA McCue K.F., Hurkman W.J., Tanaka C.K., Anderson O.D.;
RT "Starch Branching Enzymes Sbe1 and Sbe2 from Wheat (Triticum aestivum
kv, Cheyenne): Molecular Characterization, Developmental Expression,
and Homolog Assignment by Differential PCR";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF286320; AAC27624.1; -
DR InterPro; IPR001296; Glycosyltransf_1.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00534; Glycosyltransf_1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 605 AA; 66326 MW; 21120D9D6F68B891 CRC64;

```

```

Query Match 40.9%; Score 1380; DB 10; Length 605;
Best Local Similarity 51.5%; Pred. No. 3.3e-73;
Matches 281; Conservative 72; Mismatches 153; Indels 40; Gaps 9;
QY 2 LDIWVAAEVAEPWMSKYGGLGDTVGGPLPIELVKRHRVMTIAPRYDQYADAWDTSVVVDIM 61
Db 78 MNLVFGAEMAPWSKYGGLGDLVGLGPPAANGHRVMTISPRYDQYKDAWDTSVVSEIK 137

```

Db 78 NNLVEVGAEMAPWSKTGGGLDVLGGLPPAMAANGHRVMSIPRYDQKDAWDTSVVSEIK 137  
 QY 62 ----GEKVRYFHSIKGGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFAIPC 117  
 Db 138 VADEYERVRFFHCYKRGVDRVFDHPFCLEKVRGKTKKAIYGPDAIGTIDEDNQLRSLLC 197  
 QY 118 KAAIEAARVL-----PF--GP--GEDCVFVANDHWSALVPLVLLKDEYQPKGQFTKAKSVLA 169  
 Db 198 QAALAEAPRIILDNPNYFSGPYGVEDVFCNDWHITGLLACLYLKSNOSSGIYRTAKVAF 257  
 QY 170 IHNIAFQGRMWEAEAKDTKLPOAAFDKLAFCSDGYAKVYTEATPMEDEKPPLTGKTKYKI 229  
 Db 258 IHNISYQGRFSFDDFAQLNLPDRFKSSFDIDYDK-----PVVG---RKI 300  
 QY 230 NNLKGGIIAADKLVTVSPNYATEIAADAAGGVDELDTVIRAKGIEGIVNGMDIEWNPKT 289  
 Db 301 NNMKAGIIQADKLVTVSPYIAEELISGEARGCELDNIMRLTGTGIVNGMDVSEWPAKD 360  
 QY 290 KELSAPYDONSYYAGKAAKALQAEGLPVDPTAPLFAFIFGRLEQKGVDDIILAALPKI 349  
 Db 361 KFLAANYDVTTALEGGKALNKEALQAEVGLPVDKRVPLVAFIFGRLEQKGVDDVMAAIP 420  
 QY 350 LATPKVQIAILGTGAAYEKLNAIGTKYKGRAGKVVKFSAPLAHMLTAGADFMVPSRF 409  
 Db 421 LKEDYQIYLLGTGKKKFFELLSVEEKFPSKVRVVRFNAPLAHQMGAGADVLVTSRF 480  
 QY 410 EPCGLIOLHAMHYGTVPVVASTGGGLVDTVYKEGVTGFHMGALNPD--KLDEADADALAATV 467  
 Db 481 EPCGLIOLQGMRYGTPCASCSTGGGLVDTTNEKGTGFHMGHLSVDCNVVPEADVKVVTIL 540  
 QY 468 RRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLEEV-----VYKGGVATAKK 519  
 Db 541 KRAVKVVGTPAYHEMWKNCMIQDLSWSKPAKWNEDVLELGVGSEPPGVIGE--EIAPLAM 599  
 QY 520 EEIKVP 525  
 Db 600 ENVAAP 605

RESULT 14  
 Q9SQ58 PRELIMINARY; PRT; 599 AA.  
 AC Q9SQ58;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Granule-bound starch synthase GBSSI.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nakamura T., Vrinten P.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF109395; AAF14233.1;  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Glycosyltransferase; Transferrase.  
 SQ SEQUENCE 599 AA; 66045 MW; D486EFC90557F131 CRC64;

Query Match 40.9%; Score 1379; DB 10; Length 599;  
 Best Local Similarity 50.4%; Pred. No. 3.7e-73;  
 Matches 277; Conservative 78; Mismatches 147; Indels 48; Gaps 10;

QY 2 LDIYVWAEEVAPWSKTGGIGDVTGGLPIELVKGRHVMTIAPRYDOYADAWDTSVVVDI 61  
 Db 72 MPIIFVATVHPWCKTGGIGDVTGGLPIELVKGRHVMTIAPRYDOYADAWDTSVVVDI 131  
 QY 62 ----GEKVRYFHSIKGGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFAIPC 117  
 Db 132 VGDTEVVRFFHCYKRGVDRVFDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFAIPC 191

QY 118 KAAIEAARVLPF-----GP--GEDCVFVANDHWSALVPLVLLKDEYQPKGQFTKAKSVLA 169  
 Db 192 LAALAEAPRVLNNSSEYFSGPYGENVFFVANDWHITAVLPCYLKSMYKONGIYVNAKVAFC 251  
 QY 170 IHNIAFQGRMWEAEAKDTKLPOAAFDKLAFCSDGYAKVYTEATPMEDEKPPLTGKTKYKI 229  
 Db 252 IHNIAFQGRMWEAEAKDTKLPOAAFDKLAFCSDGYAKVYTEATPMEDEKPPLTGKTKYKI 294  
 QY 230 NNLKGGIIAADKLVTVSPNYATEIAADAAGGVDELDTVIRAKGIE--GIVNGMDIEWNPKT 288  
 Db 295 NNMKAGITECDVLTVSPHYVKELTSQPEKGVDELDTVIRAKGIE--GIVNGMDVVDWNPAT 354  
 QY 289 KELSAPYDONSYYAGKAAKALQAEGLPVDPTAPLFAFIFGRLEQKGVDDIILAALPK 348  
 Db 355 DKYISVKNATTVAAERALNKEALQAEVGLPVDSSIPVIVFIFGRLEQKGVDDIILAALPK 414  
 QY 349 ILATPKVQIAILGTGAAYEKLNAIGTKYKGRAGKVVKFSAPLAHMLTAGADFMVPSR 408  
 Db 415 FL--EENVQIILVIGIGKMKKEEELMLEAKYPQNAAGIENVLAHMMFAGANFIIVPSR 473  
 QY 409 FEPGGLIOLHAMHYGTVPVVASTGGGLVDTVYKEGVTGFHMGALNP--DKLDEADADALAAT 466  
 Db 474 FEPGGLIOLQGMRYGTVPIGCSSTGGGLVDTVSEGTGFHMGSFNVEFETVDPADVAASN 533  
 QY 467 RRASEVFAGGRYPPEMVANCISQDLSWSKPAQKWEGLEEVYKGG-----GVA 515  
 Db 534 VTRALKQYKTPSFHAMVQNCMAQDLSWSKPAKWN---EBALLGLGVGSGQPIEGEIEA 589  
 QY 516 TAKKEIKVP 525  
 Db 590 PLAKONVATP 599

RESULT 15  
 Q43012 PRELIMINARY; PRT; 606 AA.  
 ID Q43012  
 AC Q43012;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Starch granule-bound starch synthase (EC 2.4.1.11).  
 GN WAXY.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN--SSP. L.F. SPONTANEA;  
 RA Wang Z.Y., Zheng F.Q., Gao J.P., Wang X.Q., Wu M., Zhang J.L.,  
 RA Hong M.M.;  
 RT "Identification of two transposon-like elements in rice Wx gene."  
 RL Sci. China B 37:437-447(1994).  
 DR EMBL; X64108; CAA45472.1;  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR InterPro; IPR000794; Ketoacyl-synt.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Glycosyltransferase; Transferrase.  
 SQ SEQUENCE 606 AA; 66323 MW; E1157468A5553CE CRC64;

Query Match 40.8%; Score 1376; DB 10; Length 606;  
 Best Local Similarity 50.7%; Pred. No. 5.7e-73;  
 Matches 277; Conservative 77; Mismatches 148; Indels 44; Gaps 10;

QY 2 LDIYVWAEEVAPWSKTGGIGDVTGGLPIELVKGRHVMTIAPRYDOYADAWDTSVVVDI 60  
 Db 83 MNVYVGAEMAPWSKTGGIGDVTGGLPIELVKGRHVMTIAPRYDOYADAWDTSVVVDI 142  
 QY 61 ---MGEKVRYFHSIKGGVHRVWIDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFAIPC 117  
 Db 143 VADRYERVRFFHCYKRGVDRVFDHPFLAKVWGTGSKLYGPRSGADYLDNHRKFAIPC 202

Search completed: June 4, 2003, 14:56:01  
Job time : 37.3424 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:49:26 ; Search time 11.5821 Seconds  
(without alignments)  
1653.790 Million cell updates/sec

Title: US-09-980-771A-5

Perfect score: 3370

Sequence: 1 ALDIVVAAEVAFWPSKGTGL.....SASKTSAKPLVSAATRKSA 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*

- 1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2.6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1397.5	41.5	609	3	US-08-941-445A-7
2	1360.5	40.4	600	4	US-09-388-743-22
3	1353.5	40.2	533	3	US-08-941-445A-5
4	1342.5	39.8	616	4	US-09-388-743-14
5	1315	39.0	614	4	US-09-388-743-18
6	1314.5	39.0	615	4	US-09-388-743-2
7	902.5	26.8	641	4	US-08-836-567-10
8	882.5	26.2	649	4	US-09-192-909-2
9	881	26.1	671	4	US-09-196-390-2
10	873	25.9	801	4	US-09-388-743-26
11	871	25.8	767	4	US-08-836-567-8
12	861.5	25.6	539	3	US-08-941-445A-21
13	861.5	25.6	583	3	US-08-941-445A-13
14	858	25.5	690	4	US-09-388-743-6
15	845	25.1	558	4	US-08-836-567-6
16	841	25.0	799	4	US-09-196-390-6
17	837	24.8	698	3	US-08-941-445A-11
18	726	21.5	459	4	US-08-836-567-4
19	701.5	20.8	669	3	US-08-941-445A-9
20	580	17.2	477	1	US-07-735-065-2
21	580	17.2	477	1	US-08-469-202-12
22	580	17.2	477	2	US-08-484-434C-12
23	568.5	16.9	677	4	US-08-836-567-2
24	568.5	16.9	1197	4	US-08-836-567-12
25	568.5	16.9	1230	2	US-08-968-542C-35
26	545	16.2	735	4	US-09-115-704-2
27	530.5	15.7	533	4	US-09-388-743-10

28 504 15.0 1674 2 US-08-968-542C-12 Sequence 12, Appl  
29 240 7.1 64 2 US-08-470-720-15 Sequence 15, Appl  
30 215.5 6.4 79 2 US-08-470-720-13 Sequence 13, Appl  
31 145.5 4.3 2035 1 US-08-046-585-5 Sequence 5, Appl  
32 145.5 4.3 2035 1 US-08-393-703-5 Sequence 5, Appl  
33 145.5 4.3 2035 5 PCT-US93-11721-5 Sequence 14, Appl  
34 143.5 4.3 59 2 US-08-470-720-14 Sequence 5, Appl  
35 134.5 4.0 1222 2 US-08-682-517-15 Sequence 15, Appl  
36 134.5 4.0 1252 2 US-08-682-517-9 Sequence 9, Appl  
37 132 3.9 520 4 US-09-000-016-7 Sequence 7, Appl  
38 132 3.9 520 4 US-09-514-340-7 Sequence 7, Appl  
39 132 3.9 734 4 US-09-000-016-4 Sequence 4, Appl  
40 132 3.9 734 4 US-09-514-340-4 Sequence 4, Appl  
41 132 3.9 823 4 US-09-000-016-2 Sequence 2, Appl  
42 132 3.9 823 4 US-09-514-340-2 Sequence 2, Appl  
43 129 3.8 30 2 US-08-470-720-9 Sequence 9, Appl  
44 129 3.8 1125 4 US-09-513-783A-152 Sequence 152, App  
45 129 3.8 1610 4 US-09-513-783A-22 Sequence 22, Appl

## ALIGNMENTS

RESULT 1  
US-08-941-445A-7  
; Sequence 7, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P.  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 609 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-941-445A-7

Query Match 41.5%; Score 1397.5; DB 3; Length 609;

Best Local Similarity 51.4%; Pred. No. 1.4e-113;

Matches 281; Conservative 76; Mismatches 147; Indels 43; Gaps 9;

QY 2 LDIVVAAEVAFWPSKGTGLGVDTGGTGLPTELVKGRHVTIAPRYDOYDAWDTSVVVDI- 60

DB 83 MNVVFVGAEPWPSKGTGLGVDTGGTGLPFPAMAANGHRVMVISPRTQYKDAWDTSVVAETK 142

QY 61 ---MGEKRYFHSIKKGVHRYWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117  
DB 143 VADRYVRFFHCYKRGVDRFIDPSPLEKVGWGTGKIYGPDTGVGYKDNQMRFSLLC 202  
QY 118 KAAIEAARVL-----PFGP---GEDCVFVANDHWSALVPVLAKDEYQPKQFTKAKSVLA 169  
DB 203 QAALEAPRLNINNNPYEKGTYGEDVWFVENDWHTGPLASYLKNNYQPNGIYRNKAVFC 262  
QY 170 IHNIAFOGRMWEAEKDKLPQAAFDKLAFCSDGYAKVYTEATPMEDEKPLTGTGTYKKI 229  
DB 263 IHNISYQGRFAFEDYPELNLSEFRSSEDFIDY-----DTPVEG-----RKI 305  
QY 230 NWLKGIIAADKLVTSPNATYELAAADAAGGVDELDTVIRAKGIEGIVNGMDIEEWNPKTD 289  
DB 306 NNMKAGILEADRVLVSPYIAEELISGIARGCELDNIMELTGTIGVINGMDYSEWDPSKD 365  
QY 290 KFLSAPYDONSIVAGKAAKALQAEGLPVDPTAPLAFIAGRLBEQKGVDIILAAALPKI 349  
DB 366 KYITAKYDATTAEAKALKEALQAEGLPVDKRIPLIAFIRGLEEQKGPDMVAAAIPEL 425  
QY 350 LATPKVQIAILGTGKAAAEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRF 409  
DB 426 M-QEDVQIIVLLGTGKKEKFLKSMEEKIPGKVRVAVKKNAPLAHLMAGADVLPVPSRF 484  
QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDYTKGVTGFHMGALNPD--KLDDEADADALAATV 467  
DB 485 EPCGLIQLQGMRYGTPCASCSTGGLVDYTKGVTGFHMGRLSVDCWVWEPSPDKVAAATL 544  
QY 468 RRASEVFAGGRYPEMVANCIQSOLSWSKPAQKWEGLLEEVVYKGG-----VATK 518  
DB 545 KRAIKVVGTPAYEEMVRNMCNODLSWKGPAKNWNL--LGLGVAGSAPFIEGDEIAPLA 602  
QY 519 KEEIKVP 525  
DB 603 KENVAAP 609  
RESULT 2  
US-09-388-743-22  
; Sequence 22, Application US/09388743  
; Patent No. 6423886  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: No. 6423886el Starch Synthese polynucleotides and Their  
; FILE OF INVENTION: Use in the Production of New Starches  
; FILE REFERENCE: 1144  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 22  
; LENGTH: 600  
; TYPE: PRT  
; ORGANISM: Tulipa fosteriana  
US-09-388-743-22  
Query Match 40.4%; Score 1360.5; DB 4; Length 600;  
Best Local Similarity 49.9%; Pred. No. 2.3e-110;  
Matches 272; Conservative 77; Mismatches 157; Indels 39; Gaps 9;  
QY 2 LDIWVAAAEVAPWSKTGGLGDTVGGPLFIELVKGHRVMTIAPRDYQADAWDTSVVVDI- 60  
DB 73 MNLVFTGTGPKYSGTKGGLGDLVGLPPALAAARGHRVMTIAPRDYQADAWDINTVLEIK 132  
QY 61 MGEK---VRYFHSIKKGVHRYWIDHPWFLAKVWGTGSKLYGPRSGADYLDNHRKRALFC 117  
DB 133 VGDKMETVRFFHLKRGVDRFIDHPWFLKVGWGTGKIYGPDTGVGYKDNQMRFSLLC 192  
QY 118 KAAIEAARVLPE-----GP-GEDCVFVANDHWSALVPVLAKDEYQPKQFTKAKSVLA 169  
DB 193 QAALEAPRLNINNNKYSFGYGDVGVFIANDWHTGIPCLYKLSVYKSEGLYESAKVFC 252

QY 170 IHNIAFOGRMWEAEKDKLPQAAFDKLAFCSDGYAKVYTEATPMEDEKPLTGTGTYKKI 229  
DB 253 IHNMAVQGRFAFEDPDELLNLPDTFFKSSFOFFDGYTK-----PVKG---RKI 295  
QY 230 NWLKGIIAADKLVTSPNATYELAAADAAGGVDELDTVIRAKGIEGIVNGMDIEEWNPKTD 289  
DB 296 NNMKAGILEADRVLVSPYIAEELISGIARGCELDNIMELTGTIGVINGMDYSEWDPSKD 365  
QY 290 KFLSAPYDONSIVAGKAAKALQAEGLPVDPTAPLAFIAGRLBEQKGVDIILAAALPKI 349  
DB 356 KFITANYDATMTYAKRVNKOELQAEGLPVDPIPIVIVFVGLBEQKGSIDILAAAIPEL 415  
QY 350 LATPKVQIAILGTGKAAAEKLVNAIGTKYKGRAGVVKFSAPLAHMLTAGADFMVPSRF 409  
DB 416 M-DENVQIIVLLGTGKKEKFLKSMEEKIPGKVRVAVKKNAPLAHLMAGADVLPVPSRF 474  
QY 410 EPCGLIQLHAMHYGTVPVVASTGGLVDYTKGVTGFHMGALNPD--KLDDEADADALAATV 467  
DB 475 EPCGLIQLQGMRYGTPCASCSTGGLVDYTKGVTGFHMGRLSVDCWVWEPSPDKVAAATL 544  
QY 468 RRASEVFAGGRYPEMVANCIQSOLSWSKPAQKWEGLLEEVVYKGGVAVAKKE 520  
DB 535 KRAIKVVGTPAYEEMVRNMCNODLSWKGPAKNWNL--LGLGVAGSAPFIEGDEIAPLA 594  
QY 521 EIKVP 525  
DB 595 NVAAP 599  
RESULT 3  
US-08-941-445A-5  
; Sequence 5, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanning  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941.445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 533 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-941-445A-5

Query Match	40.2%;	Score 1353.5;	DB 3;	Length 533;
Best Local Similarity	50.4%;	Pred. No. 7.9e-110;		
Matches 275;	Conservative 79;	Mismatches 153;	Indels	Gaps 11
QY	2	LDIVMVAEAPW\$KGTGLGVDVTGGLPIELVKGRHVRMTIAPRYOQYDADWDT\$VVVDI-	60	
Db	5	MNVFVGAEAPW\$KGTGLGVDVLGGLP\$AMAANGHRV\$MVSPRYOQYDADWDT\$VVSEIK	64	
QY	61	MG---EKVYRFHSIKKGVRHVIDHPFLAKVWGKTG\$KLYGPRSGADYLDNHKRFALFC	117	
Db	65	MGDGYETVRFHCYKRGVDYRVFDHPLFLERVWGKTEEIKYGPVAGTGYDRNQLR\$LLC	124	
QY	118	KAAIEAARVL-----PF--GP-GEDEVFANDH\$ALVPVLKIDBYQPKG\$FTKAK\$VLA	169	
Db	125	QAALAEAPRL\$LN\$NPFY\$SGYGEDV\$FV\$CNDH\$TGPL\$CYL\$KSNYQ\$SHGIV\$RDAT\$FC	184	
QY	170	IHN\$AFOGRMWE\$EAFKDKL\$QA\$F\$DKL\$AF\$SDGYAKVY\$TEATP\$MEDEK\$PPLT\$GKY\$YKI	229	
Db	185	IHN\$SYQGR\$F\$AF\$Y\$PEL\$NL\$PER\$K\$SP\$DF\$IDG\$Y\$K-----P\$VEG---R\$KI	227	
QY	230	NWLKGGIIA\$DKLTV\$SNYAT\$EATA\$DAAG\$GVELD\$TVIRAKGIEGIVNGMDEEN\$PKTD	289	
Db	228	NW\$KAGI\$EAD\$RLTV\$SPY\$AEELIS\$G\$TAR\$CELD\$NIN\$RL\$TGITGIVNGMDV\$EWD\$PSD	287	
QY	290	KFL\$APYDONS\$Y\$Y\$K\$AA\$K\$AL\$Q\$EL\$GL\$PVD\$P\$APL\$F\$AF\$TGR\$LEE\$Q\$K\$GVDI\$LAAL\$PKI	349	
Db	288	KYI\$AVK\$YD\$V\$T\$AVE\$AK\$LN\$KEAL\$Q\$EVL\$G\$P\$YDR\$NIP\$IV\$AF\$TGR\$LEE\$Q\$K\$GPDV\$MAA\$IPOL	347	
QY	350	L-ATPK\$Q\$VQI\$ALGT\$G\$K\$AY\$E\$KLV\$NAIG\$PKY\$KRAK\$GVV\$K\$F\$APL\$HMLT\$ACAG\$F\$MLV\$PSR	408	
Db	348	MEMVED\$Q\$IVL\$G\$T\$G\$K\$K\$F\$ER\$ML\$S\$AE\$K\$F\$G\$K\$V\$AV\$V\$K\$F\$N\$RA\$LA\$H\$H\$MAG\$ADV\$LA\$V\$YSR	407	
QY	409	FEPCGLIQL\$HAM\$Y\$T\$VPV\$V\$AST\$GGL\$VD\$TV\$K\$E\$G\$V\$T\$G\$F\$H\$M\$G\$AL\$N\$P\$D---KLDE\$AD\$ALA\$AT	466	
Db	408	FEPCGLIQL\$G\$M\$RY\$T\$G\$T\$C\$AC\$AST\$GGL\$VD\$T\$IE\$G\$K\$T\$G\$F\$H\$M\$G\$RL\$S\$VDC\$N\$V\$VE\$AD\$V\$K\$K\$V\$AT	467	
QY	467	VRR\$ASEY\$F\$AG\$G\$RY\$E\$M\$V\$ANCIS\$QDLS\$K\$K\$PA\$OK\$W\$E\$GL\$EE---VY\$Y\$K\$G\$V-----AT\$AKK	519	
Db	468	LQRAIKV\$G\$V\$P\$AY\$E\$M\$V\$N\$C\$MIQDLS\$K\$G\$K\$AK\$N\$W\$EN\$V\$LL\$S\$G\$V\$AG\$G\$E\$F\$G\$E\$E\$T\$APL\$AK	527	
QY	520	E\$IK\$V 525		
Db	528	ENVAAP 533		

RESULT 4  
US-09-388-743-14  
; Sequence 14, Application US/09388743  
; Patent No. 6423886  
; GENERAL INFORMATION:  
; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: No. 6423886el Starch Synthase Polynucleotides and Their  
; TITLE OF INVENTION: Use in the Production of New Starches  
; FILE REFERENCE: 1114  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Canna edulis  
US-09-388-743-14

Query Match	39.8%	Score 1342.5	DB 4	Length 616
Best Local Similarity	49.5%	Pred. No. 9.1e-109		
Matches 27	Conservative	Mismatches 157	Indels 43	Gaps 10
QY	2	LDIWMVAEYAPNSKKTGGLDGVTLPIELVKGHRVMTIAPRYDQYADAWDTSVVYDI-	60	
Db	90	NNLFFVGCEYAPNSKKTGGLDGVTLRGLPMAAAIGHRVMTVVPYDQKDIWDTSVVPVELK	149	

QY	61	MGEK---	VRYSHSIKKKGVHRVWIDHPWFI	AKVWGKTSKLYGPRSGADYLDNHKRFALPC	117
Db	150	VGDKTETVRFHCHYKRGVDVRFVDFPMFTEK	YKVGKTKGKLYGFPVTGTDYADNQLRSLLC	209	
QY	118	KAATBAARVLDF-----GP-	GEDCVFVANDWHSALVPVLKDEYQPKGQFTKAKSVLA	169	
Db	210	LAALFAPRLNLNNSKYSGPYGDDVVF	IANDWHSALLPCYUKTFQSHGYINNAKVAFC	269	
QY	170	IHNIAFOGRMBEAFKDKTLKQAAFDK	LAFSDGAKVYVTEATPMBEDEKPPLTGKTKYI	229	
Db	270	IHNIAYOGRFAFSDFELNLNPNKKSSEF	DMGDYDK-----PVKG---RKI	312	
QY	230	NWLKGGIIIAADKLTVSPNYATEIARA	ADAGGVELDTVIRAKIGIEVNGMDTEWNPKTD	289	
Db	313	NWMKAGIIECDRLTVSPYIAQEIYGV	GEKVEKGNELNIRMKITCGIYNGMDTEWNPJTD	372	
QY	290	KFLSAPYDQNSVYAGKAAAKALQALG	ELGVPDPTAPLFAFICGLEQKGVDTILAAIPKI	349	
Db	373	KYISTNYDATTVLDKAPLCKEALQAC	EGLPVNKNKLVLAFVGRLLDEQKSDILAAAIPEL	432	
QY	350	LATPKVOIAIIGTKAAATEKLVNALI	GTIKYKGRAGVWVKSAPLAHMLTAGADFMVLVPSRF	409	
Db	433	LC-ENVQVIVIGTGKKKLESETLLE	EMFPDKFRAHLKFNVLAHAINAGADILVPSRF	491	
QY	410	EPCGLIQLHAMHYFTVPVYVASTG	GLVDTVTKESVTFGHMG--ALNPDPKLDEADADALAATV	467	
Db	492	EPCGLIQLQAMRYGTLPKMCSTTG	GLVDTVTKESVTFGHMGPFVSVECDKADVKQIVETV	551	
QY	468	RRASEVAGGRYPPEWANCISODLS	WSKPAQKWEGELLEEVYKGG-----VARAK	518	
Db	552	KRALUKVGTAPFVEMIKMCMODLS	WSKPAKKWOFL--LSMGAAGSEPGIDGEEIAPLA	609	
QY	519	KEEIKVP	525		
Db	610	KENVATP	616		

```

RESULT 5
US-09-388-743-18
: Sequence 18, Application US/09388743
: Patent No. 6423886
: GENERAL INFORMATION:
: APPLICANT: Singletary, George
: APPLICANT: Zhou, Lan
: TITLE OF INVENTION: NO. 6423886el Starch Synthase Polynucleotides and Their
: TITLE OF INVENTION: Use in the Production of New Starches
: FILE REFERENCE: 1144
: CURRENT APPLICATION NUMBER: US/09/388,743
: CURRENT FILING DATE: 1999-09-02
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: FastSEQ for Windows Version 3.0
: SEQ ID NO 18
: LENGTH: 614
: TYPE: PRT
: ORGANISM: Typha latifolia
US-09-388-743-18

```

[illegible]



```

QY 3 DIVVAAEAPWPKTGLGVDVGTGLPIELVKRHRVMTIAPRY-----DOYADAWDTIS 55
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
132 NILVTAEAAPYKTKGLGVDGSLPMAAARGHRVMVSPRYLNGGSPDEKYANAVDLD 191
QY 56 V--VVDIMG--EKVYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHK 111
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 VRATVHCFGDAQAEVAFHEYRAGVDVWFVDHSSVC-----RPGUP-YGDIYGA-FGDNQF 244
QY 112 RFALFCKAAITAARVLPFGP---GEDCVFVANDHMSALVPVLLKDEYQPKGQFTKAKSVL 168
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 RFTILSHAACEAPLVLPGLGTYGKCLFLANDWHAALVPLLAAKYRPGYVVKDARSIV 304
QY 169 AIHNIAPQGRWEEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGTYYK 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 AIHNIHQVEPAPVTVNLLGPPQWYGAV---EWIPTWARAHLD-----TGET--- 351
QY 229 INWLKGGITAAADKLVTSPNYATEIAADAAGGVELDTVI--RAKGIEGIVNGMDIERWNP 286
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 VNVLKGAIAVADRILTVSOGYSWEIT-PEGGYGLHELLSSRSQSVLNGITNGIDVNDWP 410
QY 287 KTDKFLSAPYDQNSVYAGKAAAEALQAEGLVPDPTAPLFAFVIGRLEEQKGVDTILAA 346
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 STDHEASHYSINDL-SGKVQCKTDLQKELGELPIRDPCLIGFTGRLDYQKGVDTILSAI 469
QY 347 PKILATPKVQJAILGTGKAAYEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLP 406
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
470 PELMQN-DVQVMLGSGEKYEDWHRHTENLFKDKFRWVGFNVPVSHRITAGCDILMP 528
QY 407 SRPECGLIQLHAMHYGTVPVVAAGTGLVDIVK-----EGV---TGPHMGALNPKLD 456
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
529 SRPECGLNQLYAMRYGTPIVHSTGGLRDTVKNDFENPYAQEGIGEGTGWTFPSLTSEKL- 587
QY 457 EADADALAATVRASEVFAAGRPE-----GVTEHKSWEGLMRMGMDRYSWENAAIQYE 627
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 8

```

US-09-192-909-2
; Sequence 2, Application US/09192909
; Patent No. 6307124
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Claus Froberg
; TITLE OF INVENTION: Nucleic acid molecules encoding soluble
; TITLE OF INVENTION: Starch synthases from maize
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o FISH & NEAVE
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02527
; FILING DATE: 16-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 19 918.2
; FILING DATE: 17-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794

```

```

; REFERENCE/DOCKET NUMBER: GFB-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-96-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-192-909-2

```

```

Query Match 26.2%; Score 882.5; DB 4; Length 649;
Best Local Similarity 38.4%; Pred. No. 1.5e-68;
Matches 203; Conservative 95; Mismatches 174; Indels 57; Gaps 15;

QY 3 DIVVAAEAPWPKTGLGVDVGTGLPIELVKRHRVMTIAPRY-----DOYADAWDTISV 56
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
142 NIVFTGEASPYAKSGGLGVDGSLPVALAARGHRVMVMPRYLNGISDKNYANAFYTEK 201
QY 57 VVDIM--GE-KVRYFHSIKKGVHRVWIDHPWFLAKVWGKTSKLYGPRSGADYLDNHKR 112
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 HIRIPCGGEHEVTFEYRDSVDVWFVDHPSY-----HRPGLYGDKTGA-FGDNQFR 254
QY 113 FALFCKAAITAARVLPFGP---GEDCVFVANDHMSALVPVLLKDEYQPKGQFTKAKSVLA 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
255 YTLCLCYAACAPLVLPGLGTYGKCLFLANDWHAALVPLLAAKYRPGYVVKDARSILV 314
QY 170 IHNIAPQGRWEEAFKDTLPQAAFDKLAISDGYAKVYVTEATPMEDEKPLTGTYYKI 229
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 IHNIHQVEPAPVTVNLLGPPQWYGAV-----VPEWARHRLDKG-----EAV 361
QY 230 NWLKGITAAADKLVTSPNYATEIAADAAGGVELDTVI--RAKGIEGIVNGMDIERWNP 287
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
362 NELKGAIVADRILTVSOGYSWEIT-PEGGYGLHELLSSRSQSVLNGITNGIDVNDWP 420
QY 288 TDKFLSAPYDQNSVYAGKAAAEALQAEGLVPDPTAPLFAFVIGRLEEQKGVDTILAA 347
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 TDKCIPCHYSVDL-SGKAKCGALQELGELPIRDPVPLIGFTGRLDYQKGVDTILQIIP 479
QY 348 KILATPKVQJAILGTGKAAYEKLVAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLP 407
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
480 DLM-REDVQVMLGSGDELEDMWRSTESIFDKFRGWGVSFVPSHRITAGCDILMPS 538
QY 408 RFPECGLIQLHAMHYGTVPVVAAGTGLVDIVK-----EGVTGPHMGALNPKLDEA 458
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
539 RFPECGLNQLYAMQYGTVPVVAAGTGLVDIVENFPEGNGEQTGWAFAPLTEN---- 594
QY 459 DADALAATVRASEVFAAGR--YPMVAVANCISQDLSWSKPAQKWEGLLE 505
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
595 ---MLWTLRTAISTYREHKSWEGMLKRGMSKDFTDWHAAYEQEYQIFQ 639

```

## RESULT 9

```

US-09-196-390-2
; Sequence 2, Application US/09196390
; Patent No. 6307125
; GENERAL INFORMATION:
; APPLICANT: Block, Martina
; APPLICANT: Loz, Horst
; APPLICANT: Lutticke, Stephanie
; APPLICANT: Walter, Lennart
; APPLICANT: Froberg, Claus
; APPLICANT: Kossmann, Jens
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES
; TITLE OF INVENTION: FROM WHEAT WHICH ARE INVOLVED IN STARCH
; TITLE OF INVENTION: SYNTHESIS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York

```

COUNTRY: United States of America  
 ZIP: 10020  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/196,390  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 196 21 588.9  
 FILING DATE: 29-MAY-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 196 36 917.7  
 FILING DATE: 11-SEP-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP97/02793  
 FILING DATE: 28-MAY-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haley, Jr., James F.  
 REGISTRATION NUMBER: 27,794  
 REFERENCE/DOCKET NUMBER: AGREVO-9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 596-9000  
 TELEFAX: (212) 596-9090  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 671 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-196-390-2

Query Match 26.1%; Score 881; DB 4; Length 671;  
 Best Local Similarity 35.2%; Pred. No. 2.2e-68;  
 Matches 237; Conservative 94; Mismatches 225; Indels 118; Gaps 22;

QY 4 IVVAAFAVAPWSKTGGLDVTGGLPTLVKRGHRYMTIAPY-----DOYADWDTSV 57  
 DB 56 IVFVTEGAPYAKSGGLDVGCSLPTALARGHRVWVMPRYLNGSSDKNYAKALYAKH 115  
 QY 58 VDIN----GBKRYFHSIKGVHRVWIDHPWFLAKVWGKTGSKLYGPRGADYLDNHRF 113  
 DB 116 IKIPCGGSHEVTFHEYRDNVDWVFDHP-----SYHRPGS-LYGNFCA-FGDNOFR 168  
 QY 114 ALFCKAAIEARVLPFGP---GEDCVFVANDHWSALVPVLLKDEYQPKGQFTAKSVLAI 170  
 DB 169 TLICYAAACEAPLILELGGYIYGQNCMEFVNDHWSALVPVLLAKIRPYGVYRDSRLVI 228  
 QY 171 HNTAFQGRMWEAFKDTKLPOAAFDKLAISDGYAKVYTEATPMEDEKPLTGKTKYKIN 230  
 DB 229 HNTAHQGVPEASTYDGLPPEWYGALEW-----VPEWARHLDKG-----EAVN 275  
 QY 231 WLKGGIIADKLVTYSPNATETATAAAGGVELDTVI--RAKGIEGVNGMDIEENPKT 288  
 DB 276 FLKGVVTDRIVTVSOGYSWEVTT-AEGQGLNELLSRKSVLNGIVNGIDINDWNPTT 334  
 QY 289 DKLSAPYDONSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILAALPK 348  
 DB 335 DKCLPHYSVDDLG-SGKAKCAELQKELGLPVRDVPDLTGIFGRDYLQKGLDILKMAIPE 393  
 QY 349 ILATPKVQIAILGTGKAAEKLVAIGTKYKRAKGVYKFSAPLAHMLTAGADPMVPSR 408  
 DB 394 LM-REDYQFVNLGSDPIFEGWMBSTESSYKDKFRGWGVFSVPVSHRITAGCDILLMP 452  
 QY 409 FEPGGLIQLHAMHYGTVPVVASVGGIVDTVK-----EGVTGFMGALNPDKLEAD 459  
 DB 453 FEPGGLNQLYAMQGTVPVVGHTGGLRDTVETENPFCAKGEEGTGWAFSPLTVDKMLWA- 511  
 QY 460 ADALAAVTRASEVFAAGRYPEMVANCISQDLSNKSQAQKWEGLLEEVYKGVATAKK 519

DB 512 -----LRTAMSTFRE-----HKPS--WEGLM-----KR 532  
 QY 520 EEIKVPVAEKIPGDLPAVSYPANTLKPVSASVEGNAAPKVGTTAPAMGAWRATTPSGP 579  
 DB 533 GMTKDHTWDHAPSSSTSRSSSGPSWNTPTSC-----RRLGRSKCESPSAL 577  
 QY 580 SPAAT---PKVTYTKPALPATAPKTAGLK-LAGEASTVSTSENGAASNGNGASASK 635  
 DB 578 KTSSSFRGPEGYPCITLRCPATVESOCACILLFAG-----SRTYDGA-----AAAVT 625  
 QY 636 TSAAKPLVSAATR 649  
 DB 626 AGGROLQFWGIRK 639  
 RESULT 10  
 US-09-388-743-26  
 ; Sequence 26, Application US/09388743  
 ; Patent No. 6423886  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Singletary, George  
 ; APPLICANT: Zhou, Lan  
 ; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their  
 ; FILE OF INVENTION: Use in the Production of New Starches  
 ; FILE REFERENCE: 1144  
 ; CURRENT APPLICATION NUMBER: US/09/388,743  
 ; CURRENT FILING DATE: 1999-09-02  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 26  
 ; LENGTH: 801  
 ; TYPE: PRT  
 ; ORGANISM: Typha latifolia  
 ; ORGANISM: Typha latifolia  
 ; US-09-388-743-26

Query Match 25.9%; Score 873; DB 4; Length 801;  
 Best Local Similarity 38.9%; Pred. No. 1.5e-67;  
 Matches 211; Conservative 79; Mismatches 166; Indels 86; Gaps 16;

QY 2 LDIVWAAEAPWSKTGGLDVTGGLPTLVKRGHRYMTIAPYDQYADWDTSV--VVD 59  
 DB 226 MNIIVAAECAPWSKTGGLDVGAGALPKALARGHRVWVAPRYNAEPQDVGKYYK 285  
 QY 60 IMGP--KVRTPHSIKGVHRVWIDHPWFLAKVWGKTGSKLY-GPRGADYLDNHRALF 116  
 DB 286 VHGDMEVTFTHAYIDGVDFVMDSPDFRHR-----GNRIYEGNR-----VDILKRMILF 335  
 QY 117 CKAIEARVLP-----FGCEDCVFVANDHWSALVPVLLKDEYQPKGQFTAKSVLAIH 171  
 DB 336 CKAIEARVLP-----FGCEDCVFVANDHWSALVPVLLKDEYQPKGQFTAKSVLAIH 394  
 QY 172 NIAFQGRMWEAFKDTKLPOAAFDKLAISDGYAKVYTEATPMEDEKPLTGKTKYKINW 231  
 DB 395 NIAHQRGPRVDDPFVGLPDHLYDLFLYD-----PVGG-----EHLNI 433  
 QY 232 LKGLIIADKLVTYSPNATETATAAAGGVELDTVI--IRGIVNGMDIEENPKTD 289  
 DB 434 FAAGLKTRADRVVTVSHGYAMELKT-SEGQGLNELLSRKSVLNGIVNGIDAKWSEED 492  
 QY 290 KFLSAP---YDONSVYAGKAAKEALQAEGLPVDPTAPLFAFGRLEEQKGVDIILA 345  
 DB 493 VHLXSDGTNYSLDTLEMGKPVCKAALQREVGLPVRNVPIATFGRDLHQKGVDLIAEA 552  
 QY 346 LPKITLATPKVQIAILGTGKAAEKLVAIGTKYKRAKGVYKFSAPLAHMLTAGADPMV 405  
 DB 553 MPWIV-SHDVQVWVMLGTGRQDLENLLRNFEQGRDKVRAWVAFSVKMAHRTITAGADILM 611  
 QY 406 PSREPCGLIQLHAMHYGTVPVVASVGGIVDTVK-----EGVTGFMGALNPDKLEADALAA 465  
 DB 612 PSREPCGLNQLYAMQGTVPVVGHTGGLRDTVETENPFCAKGEEGTGWAFSPLTVDKMLWA- 657  
 QY 466 TVRRASEVFAAGRYPEMVANCISQDLSNKSQAQKWEGLLEEVYKGVATAKK 507

Db 658 TPDRAE-----AGKLIHALNCLNTYWNKYKDSWKGLQTRGMQDLSDWNRAAQQEDVLVAA 713  
QY 508 VY 509  
Db 714 KY 715  
RESULT 11  
US-08-836-567-8  
; Sequence 8, Application US/08836567  
; Patent No. 6130367  
; GENERAL INFORMATION:  
; APPLICANT: Kossman, Jens  
; APPLICANT: Springer, Franziska  
; APPLICANT: Abel, Gernot  
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES  
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC  
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,567  
; FILING DATE: 24-JUL-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/04415  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 41 408.0  
; FILING DATE: 10-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: Agrevo-4  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 767 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-836-567-8  
Query Match 25.8%; Score 871; DB 4; Length 767;  
Best Local Similarity 39.7%; Pred. No. 2e-67;  
Matches 210; Conservative 78; Mismatches 181; Indels 60; Gaps 14;  
QY 2 LDIWVAEAPWPSKTTGLDVTGGTGLPIELVKRHRVMTIAPRYDOYADAWTSV----V 57  
Db 276 MNILVASCAPWSTKGLDVGALPKALARHGRVWVAPRIDNYPEQSGVPRKIYK 335  
QY 58 VDMGEKVYFHSIKKGVHRVWIDHPWFLAKYWGKTSKLYGPRSGADYLDNKHKRALFC 117  
Db 336 VDGQDVETVYFOAFIDGVDFEID-----SHMFRHIGNNIY-----GGNRVDILKRWLFC 386  
QY 118 KAAIEARVLP-----FGGEDCVFVANDWHSAIVPLKDEYQPKGQTTAKSVLAHIN 172  
Db 387 KAAIEVPWPCGGVYCGDG- NLVFIANDWHFALLPLFYLKAYIRONGIMNYTRSVLVIHN 445  
QY 173 IAFQGRMWEAEAFKDKLQPAADFSLAFSDGYAKVYTEATPMEDEKPLTGTGTYKINWL 232

Db 446 IAHQGRGPLEDSYVDLPPHYMDPKLYD-----PVGGEHF---NIF 484  
QY 233 KGGTIAADKLVTVSPNYATEIAADAAGVELDVTIRAKG--IEGIVMGMDTEFWNPCKDK 290  
Db 485 AAGLKTADRVVTVSHGYSWELKT-SQGGWGLHQIINENDWKLOGIVNGIDTKWNPCLDV 543  
QY 291 FLSAP-----YDQNSVYAGKAAKAEALQAEGLPVDPTAPLFAFTRGLEEKGVDIILAAAL 346  
Db 544 HLQSDGYMNSLDLFTQKPKCKAALQKELGLPVRDDVPLIGFIRGLDPQKGVLDLIAFAS 603  
QY 347 PKILATPKVQITAILGTGKAAYEKLUNVNAIGTKYKGRAGVYKFSAPLAHMLTAGADFMVLP 406  
Db 604 AMMG-QDVQLVMLGTGRRDLEQMLRQFECQHNKIRGWGFSVKTSRITAGADILLMP 662  
QY 407 SRPEPCGLIOLHAMHYGTVPVASTGGIVDTVK-----EGVTGFHMGALNPKLDEADA 460  
Db 563 SRPEPCGLNQLYAMKYGTIPVHVAGGLRDTVPQDFPFNESGLGWTFSRAEASOLIHALG 722  
QY 461 DALAATVRRASEVFPAGGRYPEMVANCISQDLSSKPAQKWEGLEEYVY 509  
Db 723 NCL-LTYREYKKSWEG-----IOTRCMTQDLSDWNRAAQNYEEVLIARAY 765  
RESULT 12  
US-08-941-445A-21  
; Sequence 21, Application US/08941445A  
; Patent No. 6107060  
; GENERAL INFORMATION:  
; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 539 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-941-445A-21  
Query Match 25.6%; Score 861.5; DB 3; Length 539;  
Best Local Similarity 40.8%; Pred. No. 7.9e-67;  
Matches 194; Conservative 81; Mismatches 154; Indels 47; Gaps 13;  
QY 4 IYVVAEAPWPSKTTGLDVTGGTGLPIELVKRHRVMTIAPRY-----DOYADAWTSV 57

Db 51 IVFVTGEASPYAKSGGLGDCVGLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 110  
QY 58 VDIM---GE-KVRYFHSIKKGVHRWIDHPWFLAKWVGKTSKLYGPRSGADYLDNHKRF 113  
Db 111 IRIPCFCGGEHEVTFHEYRDSVDWVFDHPSY-----HRPGLNYGDKFGA-FGDNQFXY 163  
QY 114 ALFCAATEAARVLPFGP---GEDCVFVANDHNSALVPVLLKDEYQPKOFTKAKSVLAI 170  
Db 164 TLLCYAACEAPLILELGGYIYQNCMFVVDNDHNSALVPVLLAARYPYGVYKDSRSILVI 223  
QY 171 HNAFQGRWEAEAFKDTKLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPLTGKTYKKIN 230  
Db 224 HNLHOGVPEASTYDGLPPEWYGALEW-----VPEWARHRLDKG-----EAVN 270  
QY 231 WLKGGIIAADKLVTSFVSPNYATEIAADAAGVLDTVI---RANGIEGIVNGMDEEWNPKT 288  
Db 271 FLKGAVVADRIVTVSKGYSWEVTT-AEGQGLNELLSRKS SVLNGVINGIDINDNPNAT 329  
QY 289 DKFLSAPYDQNSVYAKAAAEALQAEGLPVDPTAPLFAFIRGLEEQKGVDIILAAALPK 348  
Db 330 DKCIPCHYSVDL-SGKAKCKGALQELGLPIRPDVPFLIGFIRGLDYQKGDILQIIPD 388  
QY 349 ILATPKVQIAILGTGKAAAEKLVNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLPSP 408  
Db 389 LM-REDVQFVMLGSDPELDMRSTESIFKDKFRGWGFSVPVSHRITAGCDILLMPSP 447  
QY 409 FEPCGLIQLHAMHYGTVPVVAAGTGLVDTVK-----EGVTGFHMGALNPDKL 455  
Db 448 FEPCGLNQLYAMQYGTVPVVAHATGGLDRDVENENPFGEQGTGWAFAPLTENN 503

## RESULT 13

US-08-941-445A-13  
; Sequence 13, Application US/08941445A  
; Patent No. 6107060

## ; GENERAL INFORMATION:

; APPLICANT: Keeling, Peter  
; APPLICANT: Guan, Hanping  
; TITLE OF INVENTION: Starch Encapsulation  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,445A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,855  
; FILING DATE: 30-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 89-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 583 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-941-445A-13

Query Match 25.6%; Score 861.5; DB 3; Length 583;  
Best Local Similarity 40.8%; Pred. No. 8.9e-67;  
Matches 194; Conservative 81; Mismatches 154; Indels 47; Gaps 13;  
QY 4 IYVAAEAVPWSKTSGLGVDVGTGLPIELVKRHRVMTIAPRY-----DOYADAWDTSVV 57  
Db 95 IVFVTGEASPYAKSGGLGDCVGLPVALAARGHRVVMVMPRYLNGTSDKNYANAFYTEKH 154  
QY 58 VDIM---GE-KVRYFHSIKKGVHRWIDHPWFLAKWVGKTSKLYGPRSGADYLDNHKRF 113  
Db 155 IRIPCFCGGEHEVTFHEYRDSVDWVFDHPSY-----HRPGLNYGDKFGA-FGDNQFXY 207  
QY 114 ALFCAATEAARVLPFGP---GEDCVFVANDHNSALVPVLLKDEYQPKOFTKAKSVLAI 170  
Db 208 TLLCYAACEAPLILELGGYIYQNCMFVVDNDHNSALVPVLLAARYPYGVYKDSRSILVI 267  
QY 171 HNAFQGRWEAEAFKDTKLPOAFAFKLAFSDGYAKVYVTEATPMEDEKPLTGKTYKKIN 230  
Db 268 HNLHOGVPEASTYDGLPPEWYGALEW-----VPEWARHRLDKG-----EAVN 314  
QY 231 WLKGGIIAADKLVTSFVSPNYATEIAADAAGVLDTVI---RANGIEGIVNGMDEEWNPKT 288  
Db 315 FLKGAVVADRIVTVSKGYSWEVTT-AEGQGLNELLSRKS SVLNGVINGIDINDNPNAT 373  
QY 289 DKFLSAPYDQNSVYAKAAAEALQAEGLPVDPTAPLFAFIRGLEEQKGVDIILAAALPK 348  
Db 374 DKCIPCHYSVDL-SGKAKCKGALQELGLPIRPDVPFLIGFIRGLDYQKGDILQIIPD 432  
QY 349 ILATPKVQIAILGTGKAAAEKLVNAIGTKYKRAKGVVKSAPLAHMLTAGADFMVLPSP 408  
Db 433 LM-REDVQFVMLGSDPELDMRSTESIFKDKFRGWGFSVPVSHRITAGCDILLMPSP 491  
QY 409 FEPCGLIQLHAMHYGTVPVVAAGTGLVDTVK-----EGVTGFHMGALNPDKL 455  
Db 492 FEPCGLNQLYAMQYGTVPVVAHATGGLDRDVENENPFGEQGTGWAFAPLTENN 547

## RESULT 14

US-09-388-743-6  
; Sequence 6, Application US/09388743  
; Patent No. 6423886

## ; GENERAL INFORMATION:

; APPLICANT: Singletary, George  
; APPLICANT: Zhou, Lan  
; TITLE OF INVENTION: No. 6423886el Starch Synthese Polynucleotides and Their  
; TITLE OF INVENTION: Use in the Production of New Starches  
; FILE REFERENCE: 1144  
; CURRENT APPLICATION NUMBER: US/09/388,743  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Curcuma zedoaria

US-09-388-743-6

Query Match 25.5%; Score 858; DB 4; Length 690;  
Best Local Similarity 39.0%; Pred. No. 2.4e-66;  
Matches 208; Conservative 80; Mismatches 176; Indels 70; Gaps 15;

QY 2 LDIVVAAEAVPWSKTSGLGVDVGTGLPIELVKRHRVMTIAPRYDQYADAWDTSVV----- 57  
Db 199 MNILVAAECAPWSKTSGLGVDVGTGLPIELVKRHRVMTIAPRYDQYADAWDTSVV----- 258  
QY 58 VDIMEKVRVYFHSIKKGVHRWIDHPWFLAKWVGKTSKLYGPRSGADYLDNHKRFALFC 117  
Db 259 VDGQDMEIKYTHYIDSVDFVFDISP-----IFRHINDIY----GGNVDILKRWLFC 309  
QY 118 KAAIEAARVLP-----FGPSCDCVAVANDWFSALVPVLLKDEYQPKOFTKAKSVLAIHN 172  
Db 310 KAAVEVPWHPVPCGGFCYGDG-NLVFIANDWETSLLPVYLKACFRDGLMTYARGLIVHN 368





*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 14:56:21 ; Search time 20.1047 Seconds  
(without alignments)  
3342.965 Million cell updates/sec

Title: US-09-980-771a-5  
Perfect score: 3370  
Sequence: 1 ALDIWVAEFAVPSWKTGGL.....SASKTSAARPLVNAATRKSA 651

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_PEP.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB\_PEP.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCITUS\_PUBCOMB\_PEP.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_PEP.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_PEP.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_PEP.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450.5	43.0	636	9 US-10-138-075-4	Sequence 4, Appli
2	1433.5	42.5	609	9 US-10-138-075-2	Sequence 2, Appli
3	1379	40.9	599	9 US-10-138-075-5	Sequence 5, Appli
4	881	26.1	671	10 US-09-952-677-2	Sequence 2, Appli
5	841	25.0	799	10 US-09-952-677-6	Sequence 6, Appli
6	555	16.5	915	9 US-10-163-214-6	Sequence 6, Appli
7	549	16.3	914	9 US-10-163-214-12	Sequence 12, Appli
8	545	16.2	909	9 US-10-163-214-2	Sequence 2, Appli
9	537	15.9	874	9 US-10-163-214-13	Sequence 13, Appli
10	410.5	12.2	474	12 US-10-007-693-99	Sequence 99, Appli
11	402.5	11.9	476	12 US-10-007-693-69	Sequence 69, Appli
12	385.5	11.4	293	9 US-10-163-214-10	Sequence 10, Appli
13	215.5	6.1	117	10 US-09-739-438-2	Sequence 2, Appli
14	206	6.1	409	9 US-09-738-626-4738	Sequence 4738, Ap
15	189	5.6	381	9 US-09-738-626-5896	Sequence 5896, Ap
16	164	4.9	143	10 US-09-739-438-4	Sequence 4, Appli
17	155.5	4.6	191	10 US-09-924-358-22	Sequence 22, Appli
18	155.5	4.6	418	9 US-09-738-626-3951	Sequence 3951, Ap
19	136.5	4.1	385	10 US-09-767-041-18	Sequence 18, Appli

20	135	4.0	2993	9 US-09-738-626-6239	Sequence 6239, Ap
21	134.5	4.0	1222	10 US-09-137-531-15	Sequence 15, Appli
22	134.5	4.0	1252	10 US-09-137-531-9	Sequence 9, Appli
23	129	3.8	1125	9 US-10-100-957A-152	Sequence 152, Appli
24	129	3.8	1610	9 US-10-100-957A-22	Sequence 22, Appli
25	128	3.8	2910	9 US-10-124-800-2	Sequence 16, Appli
26	126	3.7	387	9 US-09-738-626-3890	Sequence 3890, Ap
27	125.5	3.7	398	10 US-09-934-899-16	Sequence 16, Appli
28	125.5	3.7	398	10 US-09-934-868-36	Sequence 36, Appli
29	124.5	3.7	1704	9 US-09-991-262-40	Sequence 40, Appli
30	124	3.7	379	9 US-09-738-626-6655	Sequence 6655, Ap
31	124	3.7	1461	9 US-10-021-955-88	Sequence 88, Appli
32	123.5	3.7	413	9 US-09-738-626-4025	Sequence 4025, Ap
33	123	3.6	1528	9 US-09-945-917-3	Sequence 3, Appli
34	123	3.6	1583	9 US-09-945-917-4	Sequence 4, Appli
35	121	3.6	339	9 US-09-847-208-116	Sequence 116, Appli
36	118.5	3.5	372	10 US-09-815-242-5277	Sequence 5277, Ap
37	118.5	3.5	384	10 US-09-815-242-12605	Sequence 12605, A
38	118	3.5	267	10 US-09-852-555-2	Sequence 2, Appli
39	118	3.5	333	9 US-09-847-208-147	Sequence 147, App
40	118	3.5	1463	9 US-09-971-536-69	Sequence 69, Appli
41	117.5	3.5	636	9 US-09-738-626-5590	Sequence 5590, Ap
42	117	3.5	3241	10 US-09-841-786-1	Sequence 1, Appli
43	117	3.5	3739	9 US-09-860-846-33	Sequence 33, Appli
44	117	3.5	3739	9 US-09-988-384B-33	Sequence 33, Appli
45	117	3.5	3739	9 US-09-836-821-33	Sequence 33, Appli

## ALIGNMENTS

### RESULT 1

US-10-138-075-4  
; Sequence 4, Application US/10138075  
; Publication No. US20030087369A1  
; GENERAL INFORMATION:  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butler, Karlene H.  
; APPLICANT: Harvell, Leslie T.  
; APPLICANT: Lightner, Jonathan E.  
; APPLICANT: Orozco, Emil M.  
; TITLE OF INVENTION: Granule-Bound Starch Synthase  
; FILE REFERENCE: BB1474 NA  
; CURRENT APPLICATION NUMBER: US/10138, 075  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/288,315  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 636  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-138-075-4

QY	2	LDIVMVAEFAVPSWKTGGLGDTYGGLPPIELVKGHRVMTIAPRYDOYADADTSSVVVDI-	60
Db	106	MTFIIGTEVAWCKTGGGLGDPALGFGHRVMTIVPRYDOYKDAWDTSSVYEVK	165
QY	61	---MGEKRYFHSIKKGVRWIDHPWFLAKYVGTGSKLYGPRSGADYLDNKKFALPC	117
Db	166	VGDRTEKVFHCYKRGVDVDFVDFHFWLEKVGKFGQKLYGTTGNDYEDNQLRESLFC	225
QY	118	KAALEAARVLPF-----GP-GEDCVFVANDWHSALVFLKDEYQPKQFTKAKSVLA	169
Db	226	QAALAEAPRVLSNSSKYSFGPYGEDIIVFVANDWHTALIPCYLKSMYQSRGIYTNARVFC	285
QY	170	INIAFOGRMWEAEKDKLPQAADFKLAFSGYAKVYVTEATPMEEDEKPTTGKTYKKI	229

Db 286 IHNIAQGRFAFADESIILNLPDQFKSFDIDGHVK-----PVVG---RKI 328  
 QY 230 NWLKGIIIAADKLTVSPNYATEIAADAAGGVVELDTVIRA-----KGIEGIVNGMDIEWN 285  
 Db 329 NWLKGIIIESWFIIVTSNYAKELVSGDPKGVDELONILIRKIDDDGLVGVINGMDVQEW 388  
 QY 286 PKTDFLSAPYDQNSVYAGKAAAKALQAEGLGPDPTAPLFAFICRLEEQKGVDIILAA 345  
 Db 389 PITDKYIAVKYDVSTVLEAKALLKEALQAEVGLPVDNRPLIFIGIRLEEQKGSILAEA 448  
 QY 346 LPKILATPKVQIATLGTGKAAAYEKLVNAIGTKYKRAKGVVWFAPLAHMLTAGADFMV 405  
 Db 449 IPQFT-KONVOLVALGTGKQKQKQLELEISYDPKARGAKFENVPVLAHMLTAGADFI 507  
 QY 406 PSREPCGLIOLHAMHYGVPVYVASTGGGLVDIVKEGVTGFHMGALNP--DKIDEADADAL 463  
 Db 508 PSREPCGLIOLQAMRYGVPVYVASTGGGLVDIVKEGVTGFHMGALNP--DKIDEADADAL 567  
 QY 464 AATVRRASEVFAAGRYPEMVANCISQDLSWSKPAQKWEGLLEVVY-----GKGG--VAT 516  
 Db 568 SKTVKRALAVYGTAPFTEIKKMCQADLSWKGPKEWEEVLLSLGVPGSEPGSDGEIAP 627  
 QY 517 AKKEEKVP 525  
 Db 628 QAKENVATP 636

## RESULT 2

US-10-138-075-2  
 ; Sequence 2, Application US/10138075  
 ; Publication No. US20030087369A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Broglie, Karen E.  
 ; APPLICANT: Butler, Karlene H.  
 ; APPLICANT: Harvelli, Leslie T.  
 ; APPLICANT: Lightner, Jonathan E.  
 ; APPLICANT: Orozco, Emil M.  
 ; TITLE OF INVENTION: Granule-Bound Starch Synthase  
 ; FILE REFERENCE: BB1474 NA  
 ; CURRENT APPLICATION NUMBER: US/10/138,075  
 ; CURRENT FILING DATE: 2002-05-02  
 ; PRIOR APPLICATION NUMBER: 60/288,315  
 ; PRIOR FILING DATE: 2001-05-03  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO. 2  
 ; LENGTH: 609  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-10-138-075-2

Query Match 42.5%; Score 1433.5; DB 9; Length 609;  
 Best Local Similarity 52.3%; Pred. No. 9.4e-94;  
 Matches 288; Conservative 75; Mismatches 139; Indels 49; Gaps 11;  
 QY 2 LDIWMAAEVAPWSKTGGLGVDVTGGLPIELVKRGHRYMTIAPRYDOYADAWDTSVVVDI- 60  
 Db 81 MTIVFVATEVHPWCKTGGLGVDVVGGLPPALAAAGHRYMTIAPRYDOYKDAWDTSVLVEV 140  
 QY 61 MG---EKVRYFHSIKKGVHRYWDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117  
 Db 141 IGTETVTRFFHCHYKRGVDVFDHPWFLAKVWGKTGAKLYGPTGTDYDNDQNLRCCLC 200  
 QY 118 KAAIEARVLFP-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKGFQTKAKSVLA 169  
 Db 201 LAALAPRVLFNNSEYFSGPYGEDVVFVANDWHTAILPCYLKSMYKPNYIYNAKVAFC 260  
 QY 170 IHNIAQGRWMEAFKDTKLPQAFDKLAFSDGYAKVYVTEATMEDEKPLPGTKYKKI 229  
 Db 261 IHNIAQGRFADFDLLNPDLSFLPSFDIDGHVK-----PVLG---RKL 303  
 QY 230 NWLKGIIIAADKLTVSPNYATEIAADAAGGVVELDTVIRAKGIE-GIVNGMDIEWNPKT 288

Db 304 NNMKAGIIBSDLVLTVSPHYVKELTSGDPKGVGLDGLVLRKPLEIGIVNGMDYEWDPST 363  
 QY 289 DKFLSAPYDQNSVYAGKAAAKALQAEGLGPDPTAPLFAFICRLEEQKGVDIILAAALPK 348  
 Db 364 DXYISAKYDATTVEARALNKEALQAEVGLPVDSSIPVIVFVGRLEEQKGSILIAAIE 423  
 QY 349 ILATPKVQIATLGTGKAAAYEKLVNAIGTKYKRAKGVVWFAPLAHMLTAGADFMVPSR 408  
 Db 424 FVG-ENVQIIVLTGKTKMEELTQLEVKYPPNARGIAKFNVPVLAHMLTAGADFIWPSR 482  
 QY 409 FEPCCGLIOLHAMHYGVPVYVASTGGGLVDIVKEGVTGFHMGALNP--DKIDEADADALAA 466  
 Db 483 FEPCCGLIOLQAMRYGVPVYVASTGGGLVDIVKEGVTGFHMGALNP--DKIDEADADALAA 542  
 QY 467 VRRASEVFAAGRYPEMVANCISQDLSWSKPAQKWEGLLEVVYKGG-----GV 514  
 Db 543 VTRALQYDTPAFHEMVQNCMAQDLSWKGPAAKKN---EVLGLIGVSGSRAGIDDAEEI 598  
 QY 515 ATAKKEEKVP 525  
 Db 599 APLAKENVATP 609

## RESULT 3

US-10-138-075-5  
 ; Sequence 5, Application US/10138075  
 ; Publication No. US20030087369A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Broglie, Karen E.  
 ; APPLICANT: Butler, Karlene H.  
 ; APPLICANT: Harvelli, Leslie T.  
 ; APPLICANT: Lightner, Jonathan E.  
 ; APPLICANT: Orozco, Emil M.  
 ; TITLE OF INVENTION: Granule-Bound Starch Synthase  
 ; FILE REFERENCE: BB1474 NA  
 ; CURRENT APPLICATION NUMBER: US/10/138,075  
 ; CURRENT FILING DATE: 2002-05-02  
 ; PRIOR APPLICATION NUMBER: 60/288,315  
 ; PRIOR FILING DATE: 2001-05-03  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 5  
 ; LENGTH: 599  
 ; TYPE: PRT  
 ; ORGANISM: Triticum aestivum  
 US-10-138-075-5

Query Match 40.9%; Score 1379; DB 9; Length 599;  
 Best Local Similarity 50.4%; Pred. No. 6.8e-90;  
 Matches 277; Conservative 78; Mismatches 147; Indels 48; Gaps 10;  
 QY 2 LDIWMAAEVAPWSKTGGLGVDVTGGLPIELVKRGHRYMTIAPRYDOYADAWDTSVVVDIM 61  
 Db 72 MPFIIVATEVHPWCKTGGLGVDVVGGLPPALAAAGHRYMTIAPRYDOYKDAWDTSVLVEVI 131  
 QY 62 ----GEKRVFHSIKKGVHRYWDHPWFLAKVWGKTGSKLYGPRSGADYLDNHRKRALFC 117  
 Db 132 VGDRTETVTRFFHCHYKRGVDVFDHPWFLAKVWGKTGSKLYGPTGTDYDNDQNLRCCLC 191  
 QY 118 KAAIEARVLFP-----GP-GEDCVFVANDWHSALVPVLLKDEYQPKGFQTKAKSVLA 169  
 Db 192 LAALAPRVLFNNSEYFSGPYGENVVFVANDWHTAILPCYLKSMYKPNYIYNAKVAFC 251  
 QY 170 IHNIAQGRWMEAFKDTKLPQAFDKLAFSDGYAKVYVTEATMEDEKPLPGTKYKKI 229  
 Db 252 IHNIAQGRFPRVDPELLNLPESPMPSFDFVDGHVK-----PVVG---RKL 294  
 QY 230 NWLKGIIIAADKLTVSPNYATEIAADAAGGVVELDTVIRAKGIE-GIVNGMDIEWNPKT 288  
 Db 295 NNMKAGIIBSDLVLTVSPHYVKELTSGPEKGVLDGLVLRKPLEIGIVNGMDVWDNPAT 354  
 QY 289 DKFLSAPYDQNSVYAGKAAAKALQAEGLGPDPTAPLFAFICRLEEQKGVDIILAAALPK 348

Db 355 DKYISVKYNATTVAEARLNKEILQAEVGLPVDSSIPVIVFIGNLEBQKSGDILLAAIPE 414  
QY 349 ILATPKVQIAILGTGAAYEKLYNAITGTYKGRAGKGVVYKFSAPLAHMLTAGADFMVLPSP 408  
Db 415 FL-EENVQIIVLTGTRKKMEELMLLEAKYPQNGRIAKENVPPLAHMMFAGANFIIVPSR 473  
QY 409 FEPGGLIQLHAMHYGVVPPVASTGGVNDTVYKEGVTGFHMGALNP--DKLDEADADALAA 466  
Db 474 FEPGGLIQLGMRIGYVPCSSSTGGVNDTVYKEGVTGFHMGALNP--DKLDEADADALAA 533  
QY 467 VRRASEVAGGRYPFEMVANCISODLSWSKPAQKWEGLLEEVYKKG-----GVA 515  
Db 534 VTRALAKYKTPSHANVQNCMAQDLSKKGPAKWK-----ERALLGLGVGSGQPGIEGEIA 589  
QY 516 TAKKEELKVP 525  
Db 590 PLAKQNVATP 599

## RESULT 4

US-09-952-677-2  
; Sequence 2, Application US/09952677  
; Patent No. US20020138876A1  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Lutticke, Stephanie  
; Walter, Lennart  
; Froberg, Claus  
; Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/952,677  
; FILING DATE: 14-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,390  
; FILING DATE: 19-No. US20020138876A1-1998  
; APPLICATION NUMBER: DE 196 21 588.9  
; FILING DATE: 29-MAY-1996  
; APPLICATION NUMBER: DE 196 36 917.7  
; FILING DATE: 11-SEP-1996  
; APPLICATION NUMBER: PCT/EP97/02793  
; FILING DATE: 28-MAY-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Haley, Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: AGREVO-9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 671 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-952-677-2

Query Match 26.1%; Score 881; DB 10; Length 671;  
Best Local Similarity 35.2%; Pred. No. 1.9e-54;  
Matches 237; Conservative 94; Mismatches 225; Indels 118; Gaps 22;  
QY 4 IVWVAARVAWSTGGTGLDVTGGLPIELVGRHVRMTIAPRY-----DQYADAWDTSVV 57  
Db 56 IVFTGTGAAPYAKSGGLDVCGLPILAAAGHRVVMVMPRYLNGSSDKNYAKALYATAKH 115  
QY 58 VDJM---GKVRVYFHSIKKGVHRVWDHDFWFLAKVWGKTSKLYGPRSGADYLDNHRF 113  
Db 116 IKPFCGSGHEVTFHEYRDNVWVFDHP-----SYHRPGS-LYGNDFGA-FGDNQF 168  
QY 114 ALFCKAAIEAARVLPFGP---GEDCVFVANDHWSALVPVLIKDEYQPKGQFTAKSLAI 170  
Db 169 TLLCYAAACEAPLILELGGYIVGQNCMFVNDHWSALVPVLLAAKRYRYGVYRDSRLVI 228  
QY 171 HNTAFQGRWEEAFKDTKLPOAAPDKLAFSDGAKYVYTEATPMEDEKPLTGTYYKKN 230  
Db 229 HNLHQGVPEASTYDGLPPEWYGALEW-----VPEWARRHALDKG-----EAVN 275  
QY 231 WLKGGIIAADKLVTVSPNYATEIAADAAGVVELDTVI--RAKGTGIVNGMDIEEWPKT 288  
Db 276 FLKGVVATDRIVTVSGISWEVIT-AEGCGQLNELLSSKSVLNGVINGIDINDWNP 334  
QY 289 DKFLSAPYDQNSYAGKAAAKEALQAEGLGVDPDTAPLFAFIGNLEBQKGVDTIILALPK 348  
Db 335 DKCLPHYSVDDL-SGKAKCKAEQLKELGVREDVPLIGFIGNLDYQKGLDKMAIPE 393  
QY 349 ILATPKVQIAILGTGAAYEKLYNAITGTYKGRAGKGVVYKFSAPLAHMLTAGADFMVLPSP 408  
Db 394 LM-REDVQFVMLGSGDPIFEGWMRSTESSYKDKFRGWGVSPVSHRITAGCDILLMP 452  
QY 409 FEPGGLIQLHAMHYGVVPPVASTGGVNDTVYKEGVTGFHMGALNPDKDEAD 459  
Db 453 FEPGGLNQLYAMQYGVVPPVHGGTGLDVTETNFPKAGEGTGWFSPDVKMLWA- 511  
QY 460 ADALATVRRASEVAGGRYPFEMVANCISODLSWSKPAQKWEGLLEEVYKKGVAATAK 519  
Db 512 -----LRTAMSTFE-----HKPS--WEGLM-----KR 532  
QY 520 EELKVPVAEKIPGDLPAVSYAPNTLKPVSASVEGNGAAAPKVGTATAMGAWRATTPSGP 579  
Db 533 GMTKHTWDHAPSTSSSSGSPSWNTPTSC-----RRGLGRKSCSPSAL 577  
QY 580 SPAAAT---PKVTYTPALPATAKPTAGLK-LAGEASTTSTENGASNGNGNGASASK 635  
Db 578 KTSSSSFERGPEGYPCITLRCPTAVESQCACILLFAG-----SRTVDGCA-----AAAVT 625  
QY 636 TSAAKPLVSAATRK 649  
Db 626 ASGGRQLQFWGIRK 639

## RESULT 5

US-09-952-677-6  
; Sequence 6, Application US/09952677  
; Patent No. US20020138876A1  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Lutticke, Stephanie  
; Walter, Lennart  
; Froberg, Claus  
; Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York



US-10-163-214-12  
 ; Sequence 12, Application US/10163214  
 ; Publication No. US20030097688A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Stephen M.  
 ; APPLICANT: Broglie, Karen E.  
 ; APPLICANT: Butler, Katherine J.  
 ; APPLICANT: Thorpe, Catherine J.  
 ; TITLE OF INVENTION: Starch Synthase Isoform V  
 ; FILE REFERENCE: BB1520 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/163,214  
 ; CURRENT FILING DATE: 2002-06-05  
 ; PRIOR APPLICATION NUMBER: 60/297,099  
 ; PRIOR FILING DATE: 2001-06-08  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 12  
 ; LENGTH: 914  
 ; TYPE: PRT  
 ; ORGANISM: Triticum aestivum  
 US-10-163-214-12

Query Match 16.3%; Score 549; DB 9; Length 914;  
 Best Local Similarity 30.8%; Pred. No. 1.1e-30;  
 Matches 169; Conservative 89; Mismatches 202; Indels 88; Gaps 22;

QY 2 LDIVMAAEVAPWSKTGGGLDVTGGPLPIELVKRGHRYMTIAPRYD-QYADAWDTSVVVDI 60  
 Db 421 LHIVHIAEAPVAKVGGGLADVISGLKALQKKGHLVEIILPKYDCMQDQVDSNKLKVLVD 480

QY 61 MGEKVRFYHSHKGVHVRWIDHPWFLAKVWGKT--GSKLY--GPR-----SGADYLDNH 110  
 Db 481 LVQS--YF-----EGNMFNKNITGIVEGLPVYFIEPQHPAMFFSRAQYYGEH 526

QY 111 ---KRFALFCAAEARVLPFGGEDC-VFVANDHSAALPVLLKDEYQPKGFTAKS 166  
 Db 527 DDFKRFYSFRAALE-----LLYQSGKKVDIIHCHDQWOTAFVAPLYWDVYANLG-FNSARI 581

QY 167 VLAHNIAPQGRWEEAFKDTKLPAQAFDKLAFSDGYAKVYVTEATPMEDEKPP-LTGKT 225  
 Db 582 CFTCHNFEYOG-----TAPARDLAWC-----GLDVEHLDRDPRDRNS 619

QY 226 YKINLWKGIIAADKLVTSVSPNYATEIAADAAGGVVELDTVIRAKGIEGVNGMDIEEWN 285  
 Db 620 HGRINAVKAVYVSNITVTSPTVAQEVRSSEGRGLQDTLVKHSKFFVGLNGIDTDTWN 679

QY 286 PTDKFLSAPYDQNSVYAGKAAKEALQAEGL-PVDPTAPLFAPIGRLEEQKGVDIILA 344  
 Db 680 PCTDRYLKVQNAKDL-QGKAANKAALREQLNLASAYSPQLVGCITRLVAQKGVHLIRH 738

QY 345 ALPKILATPKVQIATILGTGKAAEKLVAIGTKYKRA-----KGVVKFSAPLAHM 395  
 Db 739 AIYK-TAELGGQFVLLGSSP-----VPEIQREFEGADHFQNNNNIRLILKYDDALSHC 791

QY 396 LTAGADFMVPSRFEPCGLIQLHAMHYGTVPVAVASTGGLVDVT---KEGVTGFHMGAL 451  
 Db 792 IYAASDMFVPSFEPCGLTQMTAMRYGSPVYRKTYGGLNDSVDFDDEETIP---MEVRN 848

QY 452 PKLDEADADALAAVVRRASEVFAGGRYPPEMVANCISQDL-----SWSKPAQKWEGLLEE 507  
 Db 849 GFTFVKADREGLSSAMERAFNCYT--RKPEVVKQLVKQDKMTIDFSWDTASQYEDIYQKA 906

QY 508 VYKGGVA 515  
 Db 907 VARARAVA 914

RESULT 8  
 US-10-163-214-2  
 ; Sequence 2, Application US/10163214  
 ; Publication No. US20030097688A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Stephen M.

; APPLICANT: Broglie, Karen E.  
 ; APPLICANT: Butler, Katherine J.  
 ; APPLICANT: Thorpe, Catherine J.  
 ; TITLE OF INVENTION: Starch Synthase Isoform V  
 ; FILE REFERENCE: BB1520 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/163,214  
 ; CURRENT FILING DATE: 2002-06-05  
 ; PRIOR APPLICATION NUMBER: 60/297,099  
 ; PRIOR FILING DATE: 2001-06-08  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 2  
 ; LENGTH: 909  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-10-163-214-2

Query Match 16.2%; Score 545; DB 9; Length 909;  
 Best Local Similarity 29.4%; Pred. No. 2.1e-30;  
 Matches 159; Conservative 92; Mismatches 199; Indels 90; Gaps 20;

QY 2 LDIVMAAEVAPWSKTGGGLDVTGGPLPIELVKRGHRYMTIAPRYD-QYADAWDTSVVVDI 60  
 Db 416 LHIVHIAEAPVAKVGGGLADVISGLKALQKKGHLVEIILPKYDCMQHNOINNLKVLVD 475

QY 61 MGEKVRFYHSHKGVHVRWIDHPWFLAKVWGKT--GSKLY--GPRSGADYL----- 107  
 Db 476 -----VKSFEFEGNMFANKITGIVEGLPVYFIEPQHPKCFWRAQYYGEH 521

QY 108 DNHRKRFALFCAAEARVLPFGGEDC-VFVANDHSAALPVLLKDEYQPKGFTAKS 166  
 Db 522 DDFKRFYSFRAVALE-----LLYQSGKKVDIIHCHDQWOTAFVAPLYWDVYANLG-FNSARI 576

QY 167 VLAHNIAPQGRWEEAFKDTKLPAQAFDKLAFSDGYAKVYVTEATPMEDEKPP-LTGKT 226  
 Db 577 CFTCHNFEYOG-----IAPADLAYCGLDVLDRDRDR--MRDNH 615

QY 227 KINLWKGIIAADKLVTSVSPNYATEIAADAAGGVVELDTVIRAKGIEGVNGMDIEEWN 286  
 Db 616 GRINAVKAVYVSNITVTSPTVAQEVRSSEGRGLQDTLVKHSKFFVGLNGIDTDTWN 675

QY 287 PTDKFLSAPYDQNSVYAGKAAKEALQAEGL-PVDPTAPLFAPIGRLEEQKGVDIILA 345  
 Db 676 STDRLFKVQVQVANDLY-GKSANKAALKQKLAQSLASTQASQPLVGCITRLVQKGVHLIRH 734

QY 346 LPKILATPKVQIATILGTGKAAEKLVAIGTKYK--GRAKGVVKSAPLAHMLTAGADF 402  
 Db 735 IYKITELGQ-QVLLGSSPVQHQREFEGADHFQNNNNVRLLLKYDDALHMTFAASDM 793

QY 403 MLVPSRFEPCGLIQLHAMHYGTVPVAVASTGGLVDVT-----VKEGVTGFHMGAL 450  
 Db 794 FIVPSMFPEPCGLTQMTAMRYGSPVYRKTYGGLNDSVDFDDEETIPMEVRNGFTFL----- 848

QY 451 NPDKLEADADALAAVVRRASEVFAGGRYPPEMVANCISQDL-----SWSKPAQKWEGLLEE 506  
 Db 849 ---KADEQD---FGNALERAFNY--HRKPEVVKQLVKQDKMTIDFSWDTASQYEDIYQK 900

RESULT 9  
 US-10-163-214-13  
 ; Sequence 13, Application US/10163214  
 ; Publication No. US20030097688A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Stephen M.  
 ; APPLICANT: Broglie, Karen E.  
 ; APPLICANT: Butler, Katherine J.  
 ; APPLICANT: Thorpe, Catherine J.  
 ; TITLE OF INVENTION: Starch Synthase Isoform V  
 ; FILE REFERENCE: BB1520 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/163,214  
 ; CURRENT FILING DATE: 2002-06-05  
 ; PRIOR APPLICATION NUMBER: 60/297,099  
 ; PRIOR FILING DATE: 2001-06-08

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

• •

\_\_\_\_\_

\_\_\_\_\_

• •

• •

\_\_\_\_\_

[illegible]



Db 100 YSENNVVRSAFAAAA--AAYLQADPAD--IVHLHDHWYGLLAGLLKNPLN-----VHS 151  
QY 165 KSVLAHNTAFQGRMEERAFKDKLQOAFEDKLAESGYAKVYATEATPMEDEKPPITGK 224  
Db 152 KIVFTIHNFYRG-----YCSYQLAAS-----QIDDFHLSHY-----QLFRDPQTSV 194  
QY 225 TYKKINWLKGIITADKLVTSYNYATEIAADAAGGVDELTVIRAKG--TEGIVNGMDIE 282  
Db 195 L-----MKGALYCSYIITVSYVQEIINDYS-DYELHDAILARNSVFSGLINGIDE 247  
QY 283 ENPKTKDKLSAPYD-----QNSVYAGKAAAEALQAEGLPVDPTAPLFAFIRLEEQ 336  
Db 248 VNNPKTDPALAVOYDASLLSEPDVLETKKENRAVLYEKLGISD-YFPLICVISRIVEE 306  
QY 337 KG-----VDIILALPKILATPKVQIAILGTGKAAEYKLVNAIGTKYGRAGK-----VVK 387  
Db 307 KGPEFMEIILHAMESYA-----FILIGTSQ--NEVLLNEFRNLQDCLASSPNIRLILD 359  
QY 388 FSAPLAHMLTAGADEMLVPSRPEPCGLIQLHAMHYGTVPVASTGGGLVDTVKEGVTFHM 447  
Db 360 FNDPLARLYAADMICISHRACGLTQLIAMRYGTVPVLRKTGGGLADTVIPGVNGFTF 419  
QY 448 GALNPKLDEADADALAATVRRASEVFAGGRYPFEMVANCISQDLSWSKPAQKWEGLLEEV 507  
Db 420 --FDTNNE-----FRAMLSNAV-----TYRQEPDVMNLIESG 453  
QY 508 VYKGGGVATAKKEIKV 524  
Db 454 MLRASGLDAMAKHYVNL 470

## RESULT 12

US-10-163-214-10  
; Sequence 10, Application US/10163214  
; Publication No. US20030097688A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen M.  
; APPLICANT: Broglie, Karen E.  
; APPLICANT: Butier, Karlene H.  
; APPLICANT: Thorpe, Catharine J.  
; TITLE OF INVENTION: Starch Synthase Isoform V  
; FILE REFERENCE: BB1520 US NA  
; CURRENT APPLICATION NUMBER: US/10/163,214  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/297,099  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-10-163-214-10

Query Match 11.4%; Score 385.5; DB 9; Length 293;  
Best Local Similarity 33.0%; Pred. No. 9.5e-20;  
Matches 101; Conservative 59; Mismatches 115; Indels 31; Gaps 9;  
QY 228 KINWLKGIITADKLVTSYNYATEIAADAAGGVDELTVIRAKGIEGVNGMDIEENPK 287  
Db 1 RINAYKGVVYSNIVTTSPTVALEVRSEGGRLQDTLVHRSKFFIGLNGIDTDTWPC 60  
QY 288 TDKFLSAPVDQNSVYAGKAAAEALQAEGL-PVDPTAPLFAFIRLEEQKGVDTIILAL 346  
Db 61 TDRIYLVQVNAKDL-OGKAANKAALREQLNASYPSQPLVGCITRLVAQKGVHLIRHAI 119  
QY 347 PKILATPKVQIAILGTGKAAEYKLVNAIGTKYGRAGK-----KGVVKSAPLAHMLT 397  
Db 120 YK-TAELGQFVLLGSSP-----VPEIQREEGADHPQNNNNIRLILKYDDALSHCIY 172  
QY 398 AGADEMLVPSRPEPCGLIQLHAMHYGTVPVASTGGGLVDTV-----KEGVTFGFMGALNPD 453  
Db 173 AASDMFVPSIEPCGLTQMIAMRYGSPVIRKVTGGLNDSVDFDDEITP---MEVRNGF 229

QY 454 KLDEADADALAATVRRASEVFAGGRYPFEMVANCISQDL-----SWSKPAQKWEGLLEEVY 509  
Db 230 TEVKADQGLSSAMERAFNCYT--RKPEVWKQLVKQKMTIDFSWTSASQIEDIYQKAYA 287  
QY 510 KKGVA 515  
Db 288 RARAVA 293

## RESULT 13

US-09-739-438-2  
; Sequence 2, Application US/09739438  
; Patent No. US20020029394A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Beckles, Diane M.  
; APPLICANT: Thorpe, Cathy  
; TITLE OF INVENTION: Homologs of Starch Synthase DUI  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/739,438  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 60/171514  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-739-438-2

Query Match 6.4%; Score 215.5; DB 10; Length 117;  
Best Local Similarity 46.6%; Pred. No. 3.2e-08;  
Matches 41; Conservative 17; Mismatches 25; Indels 5; Gaps 1;

QY 356 QIAILGTG-----KAAEYKLVNAIGTKYGRAGVVKFSAPLAHMLTAGADFMVPSRFE 410  
Db 8 QVVLGSA PDHRIQDFTNLASKLHGEGHGRVKLCITYDEFLSHLIYAGADFILVPSMFE 67  
QY 411 PCGLIQLHAMHYGTVPVASTGGGLVDTV 438  
Db 68 PCGLTQLTAMRYGSIPIVKTGGLYDTV 95

## RESULT 14

US-09-738-626-4738  
; Sequence 4738, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIALI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4738



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 08:28:25 ; Search time 2283.88 Seconds  
(without alignments)  
16743.883 Million cell updates/sec

Title: US-09-980-771A-6  
Perfect score: 1314  
Sequence: 1 gcactgacatcgatggt.....ggcgcttggtcgacacgctc 1314

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.ov.\*
- 5: gb.pat.\*
- 6: gb.ph.\*
- 7: gb.pl.\*
- 8: gb.pr.\*
- 9: gb.ro.\*
- 10: gb.sts.\*
- 11: gb.sy.\*
- 12: gb.un.\*
- 13: gb.vi.\*
- 14: gb.ba.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.in.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pin.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1314	100.0	1314	6	AX049327	Sequence
2	1314	100.0	1593	6	AX049329	Sequence
3	1314	100.0	1593	6	AX049325	Sequence
4	1314	100.0	2124	6	AX049323	Sequence
5	1314	100.0	3117	6	AX049322	Sequence
6	1312.4	99.9	3240	8	AF026420	Chlamydom
7	511.8	38.9	5856	8	AF433156	Chlamydom
8	458.4	34.9	2289	8	AF286320	Triticum
9	456.8	34.8	1812	8	AF486514	Hordeum v
10	456.8	34.8	1827	8	AF486515	Hordeum v
11	456.8	34.8	1827	8	AF486518	Hordeum v
12	456.8	34.8	2127	8	AF163319	Triticum
13	456.8	34.8	2311	8	HVWAXYR	Barley mRNA
14	455.2	34.6	1605	8	AF250137	Triticum
15	455.2	34.6	1812	8	AF486519	Hordeum v
16	455.2	34.6	2028	8	AF113844	Triticum
17	453.6	34.5	1827	8	AF486516	Hordeum v
18	453.6	34.5	1827	8	AF486517	Hordeum v
19	432	32.9	2186	8	TAWAXYSS	Wheat waxy
20	432	32.9	2267	8	AY050174	Triticum
21	408.2	31.1	1801	8	AF113843	Triticum
22	407.6	31.0	2115	8	AB089141	Setaria i
23	400.6	30.5	2267	8	AR195560	Sequence
24	390	29.7	1937	8	AB066093	Oryza sat
25	390	29.7	1937	8	AB066094	Oryza sat
26	388.4	29.6	2542	6	AR106491	Sequence
27	388.4	29.6	2542	8	OSWX	O. sativa wa
28	386.4	29.4	2287	8	SBU23945	Sorghum bic
29	310.2	23.6	2168	8	MEGBSS	Sequence
30	306.6	23.3	2270	8	AF210699	Perilla f
31	298.6	22.7	2221	8	AB029546	Phaseolus
32	285.8	21.8	2211	8	AB071604	Ipomoea b
33	283.8	21.6	3049	8	AF026421	Chlamydom
34	281.8	21.4	2161	6	AX394246	Sequence
35	277.4	21.1	2345	8	AWA6293	Sequence
36	276.2	21.0	2124	8	IBU44126	Ipomoea bat
37	271.2	20.6	2092	8	AY094405	Arabidops
38	271.2	20.6	2115	8	AY088544	Arabidops
39	271.2	20.6	2192	8	AY123983	Arabidops
40	265	20.2	2107	8	AF097922	Astragalu
41	259.4	19.7	2035	8	PSSTARSIN	P. sativum m
42	236.8	18.0	2081	8	AF109395	Triticum
43	235.2	17.9	1698	8	AF486521	Hordeum v
44	228.4	17.4	2112	8	AY069940	Oryza sat
45	220	16.7	2100	8	PSA345045	Pisum sat

ALIGNMENTS

RESULT 1	AX049327	AX049327	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
LOCUS	AX049327	AX049327	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
DEFINITION	AX049327	AX049327	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
ACCESSION	AX049327	AX049327	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
VERSION	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
KEYWORDS	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
SOURCE	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
ORGANISM	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
REFERENCE	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
AUTHORS	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
TITLE	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001
JOURNAL	AX049327.1	GI:12226094	Sequence 6 from Patent WO0071734.	1314 bp	DNA	linear	PAT 12-JAN-2001

Db	661	CTGACGGGAAGACCTACAGAAGATCACTGGCTTAAAGGTGGCATTTATCGCCCGGCAC	720
QY	721	AAGCTGGTGACTGTGTGCCCCAACTACGCGACCGGAATATCGTGTGCGGATCGCCGCGGGGT	780
Db	721	AAGCTGGTGACTGTGTGCCCCAACTACGCGACCGGAATATCGTGTGCGGATCGCCGCGGGGT	780
QY	781	GTGGAGCTGGACACCGTCACTCGCGCCAAAGCGCATGTAGGCGCATGTGAACGGCATGGAC	840
Db	781	GTGGAGCTGGACACCGTCACTCGCGCCAAAGCGCATGTAGGCGCATGTGAACGGCATGGAC	840
QY	841	ATTGAGGAGTGGAAACCCCAAGACCGACAAGTTCCCTCTGTGGGCCCTACGACACGAACAGC	900
Db	841	ATTGAGGAGTGGAAACCCCAAGACCGACAAGTTCCCTCTGTGGGCCCTACGACACGAACAGC	900
QY	901	GTCTACGCGCGCAAGGCGCGCCGCAAGGAGGCCCTTCAAGGCGCGAGCTGGGCTGCCTGTG	960
Db	901	GTCTACGCGCGCAAGGCGCGCCGCAAGGAGGCCCTTCAAGGCGCGAGCTGGGCTGCCTGTG	960
QY	961	GACCCACACGCGCCCGCTGTTTCGCTTCAATCGCGCGCTGTGGAGGACGAGAAGGGTGTGGAC	1020
Db	961	GACCCACACGCGCCCGCTGTTTCGCTTCAATCGCGCGCTGTGGAGGACGAGAAGGGTGTGGAC	1020
QY	1021	ATCATCTGGCGCGCTGCCCAAGATCTCTGGGCCACCCCGCAAGGTGCAGATGCCATCCTG	1080
Db	1021	ATCATCTGGCGCGCTGCCCAAGATCTCTGGGCCACCCCGCAAGGTGCAGATGCCATCCTG	1080
QY	1081	GGTACCGGAAGGCGCCCTACGAGAAGCTGTGTAAGCGCATCGGCAACCAAGTACAAAGGC	1140
Db	1081	GGTACCGGAAGGCGCCCTACGAGAAGCTGTGTAAGCGCATCGGCAACCAAGTACAAAGGC	1140
QY	1141	CGCGCCAAAGGCGGTGTCTCAAGTTCTCGGCGCCCGCTGGGCGACATGCTCACGCGCGCGCC	1200
Db	1141	CGCGCCAAAGGCGGTGTCTCAAGTTCTCGGCGCCCGCTGGGCGACATGCTCACGCGCGCGCC	1200
QY	1201	GACTTCATGCTGTGTCCTCGGCTTCGAGCCCTCGGCGCTGATCCAGCTGCACGCCCATG	1260
Db	1201	GACTTCATGCTGTGTCCTCGGCTTCGAGCCCTCGGCGCTGATCCAGCTGCACGCCCATG	1260
QY	1261	CACACGGTACCGTGCCTGTGTAGCTCCACCGCGCGCGCTGGTGCACACCGCTC	1314
Db	1261	CACACGGTACCGTGCCTGTGTAGCTCCACCGCGCGCGCTGGTGCACACCGCTC	1314
RESULT 2			
AX049329		1593 bp	DNA
LOCUS		Sequence 8 from Patent WO0071734.	linear
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

BASE COUNT  
ORIGIN

290 a 539 c 510 g 254 t

YQYADAWTSVVVD:MGKVKVTFHSIKGVHVRWVDHPWFTLAKVVGKTSKSLYGPBS  
GADLDHNKRAFLCAAEAAARLFFPGEDCVFANDWHSALVPLVLLKDEYQPGKG  
TKAKSVALHNHTAQRWEEAEFKTKLPQAFKDLAFSDGYAKYVTFATPEEDEEK  
PPTGTIKKINWKGKIIAANDKLVTSPNYATEAADAAGGVLEDVTLVRAKIGIEIV  
NMDEIENPKDITKFLSDYQKQSVAGKAAKAEALQAEGLPVDPTLPAFLFAGLE  
EQGMDIILAAFLKILATPQVILGTGKAAVEKLNVATIKYKRAKGVKVSAPL  
AHMLTAGADFMLVSPFEGLQLHAMHYGTVPTVASTGGLDVDTKGEVGTGHMAL  
NPDKLADADALAATAVRASEVFAAGRYPEMWANCIQSDDLSSKSPAQKRWEGGLEEV  
YKGKGVATAKKEIKVPAEKIP"

[illegible]

Db	841	ATTGAGGAGTGGAAACCCCAAGACCCACAAGTTCCTGTCTGCGCCCTACGACCAGAACAGC	900
Qy	901	GTCATACGCCGCGCAAGGCGCGCCCAAGAGAGCCCTGCAGCGCGAGCTGGGCGCTGCCCTGTG	960
Db	901	GTCATACGCCGCGCAAGGCGCGCCCAAGAGAGCCCTGCAGCGCGAGCTGGGCGCTGCCCTGTG	960
Qy	961	GACCCACCGCGCCGCCCTGTTCGCCCTTCATCGCGCGCCTGGAGGAGCAGAAAGGGTGTGGAC	1020
Db	961	GACCCACCGCGCCGCCCTGTTCGCCCTTCATCGCGCGCCTGGAGGAGCAGAAAGGGTGTGGAC	1020
Qy	1021	ATCATCTCGCGCGCCCTGCCCAAGATCCTGCCACCCGCCCAAGGTGCAGATGCCATCCTG	1080
Db	1021	ATCATCTCGCGCGCCCTGCCCAAGATCCTGCCACCCGCCCAAGGTGCAGATGCCATCCTG	1080
Qy	1081	GGTACCGGCAAGGCGCGCTACGAGAAGCTGTGAACGCCATCGGCACCAAGTACAAGGCG	1140
Db	1081	GGTACCGGCAAGGCGCGCTACGAGAAGCTGTGAACGCCATCGGCACCAAGTACAAGGCG	1140
Qy	1141	CGCGCCAAAGGCGGTGGTCAAGTTCTCGGCGCCCTGGCGCACATGCTCACCGCGCGGCC	1200
Db	1141	CGCGCCAAAGGCGGTGGTCAAGTTCTCGGCGCCCTGGCGCACATGCTCACCGCGCGGCC	1200
Qy	1201	GACTTCATGCTGGTGCCCTCGCGCTTCGAGCCCTCGGCGCCTGATCCAGCTGCAGGCCATG	1260
Db	1201	GACTTCATGCTGGTGCCCTCGCGCTTCGAGCCCTCGGCGCCTGATCCAGCTGCAGGCCATG	1260
Qy	1261	CACHTACGGTACCGTGCGCGGTGGTAGCTCCACCGCGCGCGTGGTCGACACCGTC	1314
Db	1261	CACHTACGGTACCGTGCGCGGTGGTAGCTCCACCGCGCGCGTGGTCGACACCGTC	1314

RESULT	3
AX049325	
LOCUS	
DEFINITION	Sequence 4 from Patent WO071734.
ACCESSION	AX049325
VERSION	AX049325..1 GI:12266092
KEYWORDS	.
SOURCE	. synthetic construct. synthetic construct artificial sequences.
ORGANISM	D'Hulst,C. and Ball,S. l' (bases 1 to 1953)
REFERENCE	Starch granules containing a recombinant polypeptide of interest, method for obtaining same and uses
AUTHORS	Patient: WO 071734-A 4 30-NOV-2000.
TITLE	
JOURNAL	

FEATURES  
source  
Location/Qualifiers  
1. .1953  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="fragment de l'ADNAC complet codant la GBSII de  
Chlamydomonas reinhardtii et codant la protéine GBSII  
mature"

```

mature'
<1..>1953
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC21781.1"
/db_xref="GI:12226093"
/translation="ALDVMVAEEVAPWSKTCGLDVTGLPIELVKCKHRYMTIAPR
YDQYADAWDFSVVVLDIMGEKRVYFHSIKGVHCVWIDHPWLAKVWKGSKLYGPRS
GADYLDNHKEPALFCKAAIBAAEVLSPGCECDYFVANDHSAALPVLLIKDTQPKGQ
FTKAKSVLAHNTAFQGRWEAEAFKDTKLPLQAAFDKLAFTDAGTKATYLETAEWDEK
PMLTGKYLKLNKLKGLIAADKLVNTSNATIEAADAAGVELDVTIRAKGIEGV
NGMDIEWNNPTDKFLFASAPYDQNSVYAGKAAKEALQAEGLPVDPTVIRAFIGRL
EQGVLDLILALPKPILATPKVQILITGKAAEYKLVNAIGTKYGRAGVYKFSAPL
AHMTAGADMLVPFRFEPGQQLHAMHYTVVVAATGLVDVTKEGVTFGHMGA
NPQKLDATAADALAAIVRAKSEVFAAGRGYPEMVANGISDLSKSPQAQKAELEEV
YKGGVYATAKEIKYVPEAKITPGDLPAYSVYAPNTLKPVSAVEGNGAAAPKVGTTP
AMGAWRAITPGSPSAAAPTKVITYKPALPATAKPKTAGLKLAGEASTTSTSENGAAS
NGNGNGSASKTSAAKPLYSAPTRKSA"

```



Db 172 GCGTGGACATCGTGATGGTGTGCTGAGTGCCTTGGTCCAAAGCGGGCGGCTG 231  
QY 61 GGCATGTGACTGTGGCTTGCCTATTAGTGTGTCAAGCGCGCCACCGGTATGAGCC 120  
Db 232 GCGATGTGACTGTGGCTTGCCTATTAGTGTGTCAAGCGCGCCACCGGTATGAGCC 291  
QY 121 ATTGCCCTCGCTACGACCACTAGTGCCTGACGCGTGGGACACCTGSGTGGTGGATC 180  
Db 292 ATTGCCCTCGCTACGACCACTAGTGCCTGACGCGTGGGACACCTGSGTGGTGGATC 351  
QY 181 ATGGCGGAGAGGTCGCTACTTCCACTTCCATCAAGAGGCGTGCACCGGTGTGGATT 240  
Db 352 ATGGCGGAGAGGTCGCTACTTCCACTTCCATCAAGAGGCGTGCACCGGTGTGGATT 411  
QY 241 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCGAAGACCGGCTCCAAGCTGTACGGCCCC 300  
Db 412 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCGAAGACCGGCTCCAAGCTGTACGGCCCC 471  
QY 301 CGCTCCGGCGCTGACTACTTGGACAACCAAGCGCTTTCGCCCTGTCTGCAAGGCGGCT 360  
Db 472 CGCTCCGGCGCTGACTACTTGGACAACCAAGCGCTTTCGCCCTGTCTGCAAGGCGGCT 531  
QY 361 ATTAGAGCTGCCCGCTGCTGCCCTTCGGCCCGCGGAGGACTGCGTCTTCGTGGCCCAAC 420  
Db 532 ATTAGAGCTGCCCGCTGCTGCCCTTCGGCCCGCGGAGGACTGCGTCTTCGTGGCCCAAC 591  
QY 421 GACTGGCACTCCCGCTGGTTCGGCTCTGCTGAAGGACGAGTACCGAGGCGGAC 480  
Db 592 GACTGGCACTCCCGCTGGTTCGGCTCTGCTGAAGGACGAGTACCGAGGCGGAC 651  
QY 481 TTCACCAAGGCCAAGTCCGCTGGCTTATCCACAACATCGCTTCCAGGCGCGCATGTGG 540  
Db 652 TTCACCAAGGCCAAGTCCGCTGGCTTATCCACAACATCGCTTCCAGGCGCGCATGTGG 711  
QY 541 GAGAGGCTTTCAGGACACGAGCTGCCCGACGCGCCCTTTGACAAGCTGGCCCTTCG 600  
Db 712 GAGAGGCTTTCAGGACACGAGCTGCCCGACGCGCCCTTTGACAAGCTGGCCCTTCG 771  
QY 601 GACGGCTATCCAGGTTTACATGAGCTGAGGCGCACCCCGCATGTGAGGAGGACGAGCCCGG 660  
Db 772 GACGGCTATCCAGGTTTACATGAGCTGAGGCGCACCCCGCATGTGAGGAGGACGAGCCCGG 831  
QY 661 CTGACGGGAAGACCTACAGAAGATCACTGCTGAAGGCGGCTTATTCGCGCGCGAC 720  
Db 832 CTGACGGGAAGACCTACAGAAGATCACTGCTGAAGGCGGCTTATTCGCGCGCGAC 891  
QY 721 AAGCTGGTACTGTGCGCCCACTACGGACCGAGATCGCTCGGATGCGCGCGCGG 780  
Db 892 AAGCTGGTACTGTGCGCCCACTACGGACCGAGATCGCTCGGATGCGCGCGCGG 951  
QY 781 GTGGAGCTGGACACCGCTCATCCGCGCAAGGGGCAATTGAGGGCAATTGTAACGGCATGGAC 840  
Db 952 GTGGAGCTGGACACCGCTCATCCGCGCAAGGGGCAATTGAGGGCAATTGTAACGGCATGGAC 1011  
QY 841 ATTGAGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGAGACG 900  
Db 1012 ATTGAGAGTGAACCCCAAGACCGCAAGTTCCTGTCTGCGCCCTACGACCAAGAGACG 1071  
QY 901 GTCTAGCGCGGCAAGCGCGCCCAAGAGGCGCTTCAGGCGGAGCTGGGCGCTGCGCTGTG 960  
Db 1072 GTCTAGCGCGGCAAGCGCGCCCAAGAGGCGCTTCAGGCGGAGCTGGGCGCTGCGCTGTG 1131  
QY 961 GACCCACGCGCCCTGTGCGCTTCATCGGCGCGCTTGAGGAGGAGAGGAGGTGTGGAC 1020  
Db 1132 GACCCACGCGCCCTGTGCGCTTCATCGGCGCGCTTGAGGAGGAGAGGAGGTGTGGAC 1191  
QY 1021 ATCATCTGCGCGCCCTGCCAAGATCTCTGGCCACCCCAAGTGCAGATCGGCATCTCTG 1080  
Db 1192 ATCATCTGCGCGCCCTGCCAAGATCTCTGGCCACCCCAAGTGCAGATCGGCATCTCTG 1251  
QY 1081 GGTACCGGCAAGCGCGCTTACGAGAGCTGGTGAAGCGCATCGGCACCAAGTACAAGGGC 1140  
Db 1252 GGTACCGGCAAGCGCGCTTACGAGAGCTGGTGAAGCGCATCGGCACCAAGTACAAGGGC 1311

QY 1141 CGGCCAAGCGGTGGTCAAGTCTTCGGCGCCCTTGGCGCACATGCTACCGCCCGCGCC 1200  
Db 1312 CGGCCAAGCGGTGGTCAAGTCTTCGGCGCCCTTGGCGCACATGCTACCGCCCGCGCC 1371  
QY 1201 GACTTCATGTGGTGGCCCTCGGCTTCAGCGCTTGGAGCCCTTGGCGCTGATCCAGTGCACGCCATG 1260  
Db 1372 GACTTCATGTGGTGGCCCTCGGCTTCAGCGCTTGGAGCCCTTGGCGCTGATCCAGTGCACGCCATG 1431  
QY 1261 CACTACGGTACCGTCCGCTGTAGCTTCCACCGCGCCCTGGTGCACACCGTC 1314  
Db 1432 CACTACGGTACCGTCCGCTGTAGCTTCCACCGCGCCCTGGTGCACACCGTC 1485

RESULT 5  
AX049322  
LOCUS AX049322 3117 bp DNA linear PAT 12-JAN-2001  
DEFINITION Sequence 1 from Patent WO0071734.  
ACCESSION AX049322  
VERSION AX049322.1 GI:12226089  
KEYWORDS  
SOURCE Chlamydomonas reinhardtii.  
ORGANISM Chlamydomonas reinhardtii.  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
Chlamydomonadales; Chlamydomonas.  
REFERENCE 1 (bases 1 to 3117)  
AUTHORS D'Hulst,C. and Ball,S.  
TITLE Starch granules containing a recombinant polypeptide of interest,  
method for obtaining same and uses  
Patent: WO 0071734-A 1 30-NOV-2000;  
JOURNAL PATENT NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)  
FEATURES  
Location/Qualifiers  
source 1..3117  
/organism="Chlamydomonas reinhardtii"  
/db\_xref="taxon:3055"  
BASE COUNT 599 a 994 c 1011 g 513 t  
ORIGIN

Query Match 100.0%; Score 1314; DB 6; Length 3117;  
Best Local Similarity 100.0%; Pred. No. 5.3e-153;  
Matches 1314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGGACATCGTGATGGTGTGCTGAGTGCCTTGGTCCAAAGCGGGCGGCTG 60  
Db 186 GCGTGGACATCGTGATGGTGTGCTGAGTGCCTTGGTCCAAAGCGGGCGGCTG 245  
QY 61 GCGATGTGACTGTGGCTGCGCTATTGAGTGTGTCAAGCGCGCCACCGGTATGAGCC 120  
Db 246 GCGATGTGACTGTGGCTGCGCTATTGAGTGTGTCAAGCGCGCCACCGGTATGAGCC 305  
QY 121 ATTGCCCTCGCTACGACCACTAGTGCCTGACCGCTGGGACACCTCGGTGGTGGATC 180  
Db 306 ATTGCCCTCGCTACGACCACTAGTGCCTGACCGCTGGGACACCTCGGTGGTGGATC 365  
QY 181 ATGGCGGAGAGTGCCTACTTCCACTTCCATCAAGAGGCGGTGCACCGGTGTGGATT 240  
Db 366 ATGGCGGAGAGTGCCTACTTCCACTTCCATCAAGAGGCGGTGCACCGGTGTGGATT 425  
QY 241 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCGAAGACCGGCTCCAAGCTGTACGGCCCC 300  
Db 426 GACACCCCTGGTTCCTGGCCCAAGTCTGGGCGAAGACCGGCTCCAAGCTGTACGGCCCC 485  
QY 301 CGCTCCGGCGCTGACTACTTGGACAACCAAGCGCTTTCGCCCTGTCTGCAAGGCGGCT 360  
Db 486 CGCTCCGGCGCTGACTACTTGGACAACCAAGCGCTTTCGCCCTGTCTGCAAGGCGGCT 545  
QY 361 ATTAGAGCTGCCCGCTGCTGCCCTTGGCCCGCGGAGGACTGCGTCTTCGTGGCCAAAC 420  
Db 546 ATTAGAGCTGCCCGCTGCTGCCCTTGGCCCGCGGAGGACTGCGTCTTCGTGGCCAAAC 605.  
QY 421 GACTGGCACTCCCGCTTGGTGGCGCTTCTGCTGAAGGACGAGTACCGACCAAGGCGCAG 480  
Db 606 GACTGGCACTCCCGCTTGGTGGCGCTTCTGCTGAAGGACGAGTACCGACCAAGGCGCAG 665.





309	QY	Db	GGCTGGGACATCGTGATGGTTGTCTGAGTGCCTTGGTCCAAAGACGGCGCGCGTGG	360
61	QY	Db	GGCGATGTGACTGTGGCCCTGCCTATTAGCTGCTCAAGCGGGCCACCGCTCATGACC	120
369	QY	Db	GGCGATGTGACTGTGGCCCTGCCTATTAGCTGCTCAAGCGGGCCACCGCTCATGACC	428
121	QY	Db	ATTGGCCCTGCTACGACCAAGTACGCTAGCCCTGGGACACCTTCGGTGTCTGGACATC	180
429	QY	Db	ATTGGCCCTGCTACGACCAAGTACGCTAGCCCTGGGACACCTTCGGTGTCTGGACATC	488
181	QY	Db	ATGGCGGAGAGGTCCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATT	240
489	QY	Db	ATGGCGGAGAGGTCCCGCTACTTCCACTCCATCAAGAGGGCGTGCACCGCTGTGGATT	548
241	QY	Db	GACCACCCCTGTTTCCTGGCCAAAGTCTGGGGCAAGACCGCTTCCAAGTGTACGGCCCC	300
549	QY	Db	GACCACCCCTGTTTCCTGGCCAAAGTCTGGGGCAAGACCGCTTCCAAGTGTACGGCCCC	608
301	QY	Db	CGCTCCGGCGGTGACTACCTTGGACAAACCACAAGCGCTTCGCCCTGTTCGAAGCGCGT	368
609	QY	Db	CGCTCCGGCGGTGACTACCTTGGACAAACCACAAGCGCTTCGCCCTGTTCGAAGCGCGT	668
361	QY	Db	ATTGAGGCTCCCGCGGTGCTCCCTTCGCGCCGGCGGAGGACTGCGTCTTCGTGSCCAAC	420
669	QY	Db	ATTGAGGCTCCCGCGGTGCTCCCTTCGCGCCGGCGGAGGACTGCGTCTTCGTGSCCAAC	728
421	QY	Db	GACTGGCACTCCCGCGGTGCTCCCTTCGCGCCGGCGGAGGACTGCGTCTTCGTGSCCAAC	480
729	QY	Db	GACTGGCACTCCCGCGGTGCTCCCTTCGCGCCGGCGGAGGACTGCGTCTTCGTGSCCAAC	788
481	QY	Db	TTTACCAAGGCCAAGTTCGGTGTGCGGTATCCACAACATCGCTTCCAGGGCGCGATGTGG	540
789	QY	Db	TTTACCAAGGCCAAGTTCGGTGTGCGGTATCCACAACATCGCTTCCAGGGCGCGATGTGG	848
541	QY	Db	GAGGAGGCTTTCAAGGACACGAGTGCCTCCAGGCGCGCTTTGACAAGCTGGCTTCCTCG	600
849	QY	Db	GAGGAGGCTTTCAAGGACACGAGTGCCTCCAGGCGCGCTTTGACAAGCTGGCTTCCTCG	908
601	QY	Db	GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCGCATGTGAGGAGGACGAGAAGCCCGG	660
909	QY	Db	GACGGCTATGCCAAGGTTTACACTGAGGCCACCCCGCATGTGAGGAGGACGAGAAGCCCGG	968
661	QY	Db	CTGACGGGAAGACTTACAAGAGATCAACTGGCTGAAGGGTGGCAATTATCCCGCGCGAC	720
969	QY	Db	CTGACGGGAAGACTTACAAGAGATCAACTGGCTGAAGGGTGGCAATTATCCCGCGCGAC	1028
721	QY	Db	AAGCTGTGACTGTGTCGCCCAACTACGCGACCGAGATCGCTCCGAGTCCGCGCGCGGT	780
1029	QY	Db	AAGCTGTGACTGTGTCGCCCAACTACGCGACCGAGATCGCTCCGAGTCCGCGCGCGGT	1088
781	QY	Db	GTGGAGCTGGACACCGTCAATCCGCGCCCAAGGSCATTGAGGSCATTGTGAAGGGCATGGAC	840
1089	QY	Db	GTGGAGCTGGACACCGTCAATCCGCGCCCAAGGSCATTGAGGSCATTGTGAAGGGCATGGAC	1148
841	QY	Db	ATTGAGGAGTGGACCCCAAGACCCACAAGTTCCTGTCGTGCGCCCTAGCACAGAACAGC	900
1149	QY	Db	ATTGAGGAGTGGACCCCAAGACCCACAAGTTCCTGTCGTGCGCCCTAGCACAGAACAGC	1208
901	QY	Db	GTCTAGCCCGGCAAGGCGCCGCCCAAGGAGGCGCTTGCAGGCGGAGTGTGGAGTGTGGAG	960
1209	QY	Db	GTCTAGCCCGGCAAGGCGCCGCCCAAGGAGGCGCTTGCAGGCGGAGTGTGGAGTGTGGAG	1268
961	QY	Db	GACCCACCGCCCGCCCTGTTCCGCTTTCATCGCGCGCCCTGGAGGAGCAGAAGGGTGTGGAC	1020
1269	QY	Db	GACCCACCGCCCGCCCTGTTCCGCTTTCATCGCGCGCCCTGGAGGAGCAGAAGGGTGTGGAC	1328
1021	QY	Db	ATCATCTGGCGCCCTGCCAAGATCTTGGCCACCCCGCAAGGTGCAAGATCGCCATCTCTG	1080
1329	QY	Db	ATCATCTGGCGCCCTGCCAAGATCTTGGCCACCCCGCAAGGTGCAAGATCGCCATCTCTG	1388
1081	QY	Db	GGTACCGGCAAGGCGCCCTTACGAGAAGCTGTGTAACGCCCATCGCACCAAGTACAGGGC	1140
1389	QY	Db	GGTACCGGCAAGGCGCCCTTACGAGAAGCTGTGTAACGCCCATCGCACCAAGTACAGGGC	1448

Qy	1141	CGGCGCAAGGGCGTGGTCAAGTTCTCGGCGGCCCTTCGGCGCATAGTCTCACCGCCGGGCGC	1200
Db	1449	CGGCGCAAGGGCGTGGTCAAGTTCTCGGCGGCCCTTCGGCGCATAGTCTCACCGCCGGGCGC	1508
Qy	1201	GACITTCATGCTGGTGCCCTTCGCGCTTCAGAGCCCTTCGCGCCCTGAHCCAGTGCACGCCCATG	1260
Db	1509	GACTTCATGCTGGTGCCCTTCGCGCTTCAGAGCCCTTCGCGCCCTGAHCCAGTGCACGCCCATG	1568
Qy	1261	CACTACGGTACCGTGGCCGTGGTAGCTCCACCAGCGCGCCCTGGTCGACACCGTC	1314
Db	1569	CACTACGGTACCGTGGCCGTGGTAGCTCCACCAGCGCGCCCTGGTCGACACCGTC	1622
RESULT 7			
AF433156		5856 bp DNA linear PLN 23-AUG-2002	
LOCUS		Chlamydomonas reinhardtii granule-bound starch synthase I (STA2)	
DEFINITION		gene, complete cds.	
ACCESSION		AF433156	
VERSION		AF433156.1 GI:16755882	
KEYWORDS			
SOURCE		Chlamydomonas reinhardtii.	
ORGANISM		Chlamydomonas reinhardtii	
REFERENCE		Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;	
AUTHORS		Chlamydomonadaceae; Chlamydomonas.	
TITLE		Wattebled,F., Buleon,A., Bouchet,B., Ral,J.P., Lienard,L., Delvalle,D., Binderup,K., Dauvillee,D., Ball,S. and D'Hulst,C. Granule-bound starch synthase I: a major enzyme involved in the biogenesis of B-crystallites in starch granules	
JOURNAL		Rur. J. Biochem. 269 (15), 3810-3820 (2002)	
PUBMED		12153578	
REFERENCES		2 (bases 1 to 5856) Wattebled,F., Ball,S.G. and D'Hulst,C. Direct Submission	
AUTHORS		Submitted (15-OCT-2001) UMR576 du CNRS, Laboratoire de Chimie Biologique, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq CEDEX 59655, France	
TITLE		Location/Qualifiers	
JOURNAL		1. .5856 /organism="Chlamydomonas reinhardtii"	
PUBMED		/db_xref="taxon:3055"	
AUTHORS		1343. .>4911	
TITLE		/gene="STA2"	
JOURNAL		join(1343..1781,2151..2340,2484..2618,2815..2973,	
PUBMED		3186..3373,3582..4117,4295..>4911)	
REFERENCES		/gene="STA2"	
AUTHORS		/product="granule-bound starch synthase I"	
TITLE		join(1480..1781,2151..2340,2484..2618,2815..2973,	
JOURNAL		3186..3373,3582..4117,4295..>4911)	
PUBMED		/gene="STA2"	
FEATURES		/note="similar to the sequence deposited in GenBank	
SOURCE		Accession Number AF026420; GBSST"	
GENE		/codon_start=1	
mRNA		/product="granule-bound starch synthase I"	
CDS		/proteio_id="AAL28128.1"	
		/db_xref="GI:16755883"	
		/translation="MAVASTSPSSARPPIVINAASGVYKKTANOLLRELARGSAKST SRSYATVGATGACALDVMVAEAAPWSKVTGSLGDVTGGLPELVKGHRHYMTAPRY DOADAWTSVVVDIMGEKVYFHSIKGVHVRIWDHPFLAKVWGKTGSKLYGRSG ADYLNDHARFALFCKAALEAARLVPFGEDCVFANDWHSLAVPLLKDYEYQKGF RATGSVLAIHNIAFOGRWEAEKDTKLPOAFDKLAFSDGAKVYTEATEPMEDEKP PLTKGYKIINMLKGIITDAOKLVIVSNYATEADAAGVELDVIRAKIGIENV KMDIEWNKPTKLSFDGNSVYAGKAKEALOELGPLDPDTAPLPFAFIGRLE QGVDIILAAPLKILLTPKVOITAILGTGAAYEKLVNAIGTKYKRAKGVYFSAPLA HMLTAGDFMLVPSEFPCEGGLIQHLHAMHYGVYPVASTGGVLIVDKGVTGFHMNALN PDKLDADALAAATVRRAESVFAGGRYPENVANCISODLSWSKPAQRKEGILLEEVY GKGVDATAKKEIKVPVAEKIPGDLPAVSYPANTLIKPVSVSEGNAGAAPKVGTTAPA MGWRATTGPSGPAAATPKVITYTKPALPATAPAKTPKTAGLKAGEASTTSTSENCAASN GNGNGASASKTSAARKPLVSAATRKA"	
BASE COUNT		1073 a 1816 c 1852 g 1115 t	
ORIGIN			







```

Db 842 CCACAGAGTTCAGTCTCTTCGACTTCATTGACGGCTACGACAAG----- 888
QY 629 CCACCCCATGAGGAGGACGAGAGCCCGCTGACGGGAAGACCTACAGAAGATCA 688
Db 889 -----CCGCTGAGGGGCGC-----AAGATCA 910
QY 689 ACTGGCTGAAGGGTGGCATTTATCGCCCGCAGACAGCTGGTGAAGTCTGTCGCCCACTACG 748
Db 911 ACTGGATGAAGCGCGGATCTCGAGGCGCAGAGGCTGACGGTGAGCCCTACTACG 970
QY 749 CGACCGAGATCGTCCGATCGCCGCGCGGTGTGGAGCTGGACACCGCTCATCGCGCCA 808
Db 971 CTGAGGAGCTCATCTGCGCAAGCCAGGGCTGCGAGCTCGACAACTATGCGCTCA 1030
QY 809 AGGCATTGAGGGCATTTGTAACGCGCATGGACATGAGGATGGAACCCCAAGACCGACA 868
Db 1031 CGGGATCACCGGAATGTCACCGCATGGACGTCAGTGAGTGGAGCCCAAGGACA 1090
QY 869 AGTTCTGTCTGCGCCCTACGACAGAACAGCGTCTAGCCGCGCAAGGCGCGCCAGG 928
Db 1091 AGTTCTCGCGCTCAACTACGACATACCAACCGCTTGGAGGCCAAGCGCTGACAAAG 1150
QY 929 AGGCCCTGACGGCGAGCTGGCTGCTGTGGACCCGACCGCCCGCTGTTCGCCCTCA 988
Db 1151 AGGCATGCGAGGCGAGGTGGGCTGCGGTGGAGAGAGGTGCGGTGCGCTTCA 1210
QY 989 TCGGCCGCTGGAGGAGCAGAGGCTGTGGACATCATCTTGGCGCGCTGCCCAAGATCC 1048
Db 1211 TCGCAGGCTGGAGGAGCAGAGGCGCGGACGTGATGTCGCGCCATCCCGGAGATCC 1270
QY 1049 TGGCCACCCCAAGTGCAGATGCCATCTTGGTACCGGCAAGCGCGCTTACGAGAAGC 1108
Db 1271 TGAAGGAGGAGGAGCTCCAGATCATCTCTTGGCACCAGGGAAGAAGTTGAGAAGC 1330
QY 1109 TGTGTAAGCGCATCGGACCAAGTACAGGCGCGCGCCAGAGGCGTGGTCAAGTTCTCG 1168
Db 1331 TGCTCAAGAGCATGGAGGAGAGTTCCCGGGCAAGGTGAGGGCGGTGCTCAGTTCAAG 1390
QY 1169 CGCCCTGGCGGCATGCTCACGCGCGCGGCTGATCTGCTGGTGGTGGCTGCGGCTTCG 1228
Db 1391 CGCGCTAGCTACCAAGATGAGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
QY 1229 AGCCCTGGCGGCTGATCCAGTGCAGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1288
Db 1451 AGCCCTGGCGGCTGATCCAGTGCAGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1510
QY 1289 CCACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Db 1511 CCACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536

RESULT 11
AF486518
LOCUS
DEFINITION
Hordium vulgare cultivar SB 85750 granule bound starch synthase I
mRNA, complete cds.
ACCESSION
AF486518
VERSION
AF486518.1 GI:21667435
KEYWORDS
SOURCE
Hordium vulgare.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordium.
REFERENCE
1 (bases 1 to 1827)
AUTHORS
Patron,N., Smith,A., Fahy,B., Hylton,C., Naldrett,M., Rossnagel,B.
and Denyer,K.
TITLE
A mutation in the 5' non-coding region of the barley GBSST gene
alters its temporal and spatial expression and reduces GBSST
activity and amylose content in the endosperm
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1827)
AUTHORS
Patron,N.

```

```

Direct Submission
Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre,
Norwich Science Park, Norwich, Norfolk NR4 7UH, UK
FEATURES
Location/Qualifiers
source
1..1827
/organism="Hordium vulgare"
/cultivar="SB 85750"
/db_xref="taxon:4513"
1..1827
/feature="alpha glycosyl transferase"
/codon_start=1
/product="granule bound starch synthase I"
/protein_id="AA074051.1"
/db_xref="GI:21667436"
/translation="MAALATLSOLATSGTGLVGTDRSAPMFRHAGFQGLRPRKPADGT
FGMTVGASAPKOSRAHRCNRECLSVVVRATGSGMNLVFGAMAPWFKYKRGVDV
LGLPPAAMANGHEVWVVSRYDQYKDAWDSVISEIKVADEYERVFHCYKRGVDV
VFDHPWLEKVRGKTEKIYDPAGDIDEDNQQRFSLLQAALEAIPRLNNPYP
QNLNDRPKSDFIDYDKPVEGRKINMKAGLQADKFLAVNYDITVSPYAEELISGEARG
CELDNIMRLTGTIVGMDVSEWDPDKDFLAVNYDITVSPYAEELISGEARG
PVRKVPVLAIFIGRLBOKGPDVMIAPILKEEDVQIILGLTGKPKFLKSMEE
KFGKVRVAVRENAPLAHOMAGADLLAVTSREPCSLIQLOHRYGTPCVCASTGGL
VDRIYEGKTFHMRGLSVDCNVVPAVKVATLTKRAVKVGTPTAYOEMVKNQMIOD
LSKRGPAKNWEDVILLEGVSGEPGIVGEIAPLAVENVAAP"
BASE COUNT 364 a 567 c 602 g 294 t
ORIGIN
Query Match 34.8%; Score 456.8; DB 8; Length 1827;
Best Local Similarity 62.2%; Pred. No. 1.8e-47;
Matches 837; Conservative 0; Mismatches 422; Indels 87; Gaps 4;
QY 5 TGGACATCGTGTGATGTTGCTGCTGAGTGGCGGCTTGGTCCAAAGACGGCGGCTGGGCG 64
Db 242 TGAACCTCGTTCGTGCGCGCGAGATGGCGCTTGAGCAAGACCGCGGCTTGGCG 301
QY 65 ATGTGACTGTGGCTGCTGCTATTTAGCTGTTCAAGCGCGCCACCGCGCTCATACCATG 124
Db 302 AGTCTCTGGCGGCTTCCACAGCCATGGCGGCAACAGGTGATGTCGTCT 361
QY 125 CCCCTGCTAGCAGCAGTACGCTCACGCTGGGACACCTCGTGGTGTGTCGATCATG 184
Db 362 CCCCGCGTACATCATGATACAGGCGCTGGACACAGCGCTCATCTCCAGATCAAG 421
QY 185 -----GGGAAAGTCCGCTACTTCCACTCATCAAGAAAGGCGGTGACCGCG 232
Db 422 TCGCTGACGAGTACGAGAGGTGAGGTTCCTCCACTGCTACAGCGGAGTGGACCGG 481
QY 233 TGTGATGACCAACCGCTGTTCTGCGCAAGGTCTGGGCAAGACCGGCTTCCAGCTGT 292
Db 482 TGTTCATCGACACCGCTGTTCTGAGAAAGGTTCGCGGGCAAGACCAAGAGAAGATCT 541
QY 293 ACGGCCCGCGCTCGCGGCTGACTTACCTGGACACCAACAGCGCTTCCGCTGTTCGCA 352
Db 542 ACGGCCCGCGCGGCGGCGGCTGACGAGTACAGGACACCAAGCGGCTTCCGCTTCTG 601
QY 353 AGCGCGCTATTTAGGCTGCGCGGCTGCTGCGCTTCCGCGC----- 392
Db 602 AGGACGCTCGAGGCAACCGAGGATCCTCAACCTCAACACACACCCCTACTTCTGCTC 661
QY 393 ----CGCGGAGGACTGCTTCTGTCGCCAAGCTGCGACTCGGCTGCTGCTGCTGCTGCTC 448
Db 662 CTTACGGGAAGACGTGTTGTTGTCGAACGCTGCGACACGCGGCTTCTGCGCTGCT 721
QY 449 TGTGAAGGACGAGTACCGACCAAGGCGAGTTCACCAAGGCAAGTGGTGTGCTGCTA 508
Db 722 ACTCAAGACCACTACAGTCCCAATGGCAATACAGGACGCGCAAGGTGGCGCTTCTGCA 781
QY 509 TCACAACTGCGCTTCCAGGCGGCTGTTGGAGGAGGCTTTCAAGGACACCAAGCTGCG 568
Db 782 TCCACAACTGCTACGAGGCGGCTTCTTCCCTTCCAGGCTTTCGCGAGCTCAACTGCG 841
QY 569 CCCAGCGCGCTTTGACAAAGCTGGCGCTTCTCGAGCGGCTATGCCAAGGTTTACACTGAGG 628

```

```

842 CGACAGGTTCAAGTCGCTTCGACTTCATTGACGGCTACGACAAG----- 888
QY 629 CCACCCCCATGAGGAGGACGAGAACGCCGCCCTGACGGGAAGACCTACAGAGATCA 688
Db 889 -----CCCGTGGAGGGCGC-----AAGATCA 910
QY 689 ACTGGCTGAAGGTGCGATTATCGCGCGGCGGCAAGCTGTGACTGTGCGCCCACTACG 748
Db 911 ACTGGAAGGCGCGGATCTGCGAGCGGCGGCAAGGTGCTGACGGTGAGCCCTACTACG 970
QY 749 CGACCCAGATCGTGCCTGCGGCGGCGGCTGTGAGGTGACACCCGCTCATCCGCCCA 808
Db 971 CTGAGAGCTATCTCTGCGAAGCCAGGCGCTGCGAGCTCGACACAATCATCGCGCTCA 1030
QY 809 AGGGCAITGAGGCGATGTGAACGGCATGACATTTGAGAGTGGACCCCAAGACCGACA 868
Db 1031 CGGGGATCACCGGATCGTCAACGGCATGACGCTAGTGAAGGACCCCAAGGACA 1090
QY 869 AGTTCGTCTGCGGCTACGACCAAGACAGCGTCTACGCGGCAAGGCGCGCCAAAG 928
Db 1091 AGTTCCTCGGCTCACTACGATCACACCGCTTGGAGGCCAGGCGCTGAACAAG 1150
QY 929 AGCCCTCGAGGCCGAGCTGGGCGCTGCTGTGGACCCCAAGCGCCCTGTCGCTTCA 988
Db 1151 AGCACTGAGGCGGAGTGGGCGTCCGCTGGACAGGAAGTGGCGCTGGTGGCGCTTCA 1210
QY 989 TCGGCGGCTGAGGAGGACGAGGCTGTGACATCATCTCGCGCGGCGCTGCCCAAGTCC 1048
Db 1211 TCGGAGGCTGAGGAGGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1270
QY 1049 TGGCCACCCCAAGGTGACATCGGCTGCTGCGGTACCGGCAAGGCGGCGGCTACGAGAAGC 1108
Db 1271 TGAAGGAGGAGCGTCCAGCATCTCTCTTGGCACCGGGAAGAGAACTTTGAGAAGC 1330
QY 1109 TGTGAACGCCATCGGACCAAGTACAAGGCGCGGCGGCAAGGCGGCTGCTCAAGTTCTCGG 1168
Db 1331 TGTCAAGAGCATGGAGGAGGATTCGCGGCGGCAAGGTGAGGCGGCTGCTCAAGTTCAAGC 1390
QY 1169 CGCCCTGCGGCGCATGCTCAACGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1228
Db 1391 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1450
QY 1229 AGCCCTGCGGCGTATCCAGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1288
Db 1451 AGCCCTGCGGCGTATCCAGCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1510
QY 1289 CCACCGCGGCGTGTGCGACCGGTC 1314
Db 1511 CCACCGCGGCGTGTGCGACCGATC 1536

RESULT 12
AF163319 LOCUS AF163319 2127 bp mRNA linear PLN 14-FEB-2000
DEFINITION Triticum aestivum granule-bound starch synthase I (Wx-D1) mRNA,
Wx-D1a allele, complete cds.
ACCESSION AF163319
VERSION AF163319.1 GI:6969977
KEYWORDS Triticum aestivum.
SOURCE Triticum aestivum
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 2127)
Vrinten, P.L. and Nakamura, T.
Wheat granule-bound starch synthase I and II are encoded by
separate genes that are expressed in different tissues
Plant Physiol. 122 (1), 255-264 (2000)
20098733
MEDLINE
PUBMED 10631269
REFERENCE 2 (bases 1 to 2127)

```

```

Vrinten, P.L. and Nakamura, T.
Direct Submission
Submitted (28-JUN-1999) Crop Breeding, Tohoku National Agricultural
Experiment Station, Akihiro 4, Morioka, Iwate 020-0198, Japan
FEATURES
source
    I. .2127
    /organism="Triticum aestivum"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /chromosome="7p"
    /dev_stage="immature seed"
gene
    I. .2127
    /gene="Wx-D1"
    /allele="Wx-D1a"
    30..1844
CDS
    /gene="Wx-D1"
    /note="Waxy protein; GBSSI"
    /codon_start=1
    /product="granule-bound starch synthase I"
    /protein_id="AAF34135.1"
    /db_xref="GI:6969978"
    /translation="MAALVTSQATSGTGLTIDFRFRAGFGQVRPRSPADALGMRT
    VGASAAPTOSRKAHRTTRCLSMVVRATSGGMNLFVGAEMAPFVKTKYKRGVDRVFD
    PMAANGHRVWVISPRIQYDQDAMTSVSEIKVVDKYKRVKRYKRGVDRVFD
    HPCFLKVRGKTKKIKYGPDACTDIEDNQRFSLQCAALEVRIILNLDNNPVSQPY
    GEDVFEVNDWHTGLLACYLKSNTQNGIIRAKAVAFCTHNSISYGRFSFDDFAQLNL
    PDRKSFDFIDYDKPVEGRKINMKAGILOADKVLTVSPYAEELISGEARGELD
    NMLRTGITGVNGMDVSEWDPDKELAVNYDITTALEKALKNEALQAEPLPDR
    KVLVAFIGRIPEOKGPDVMAIAPELKEEDVQIVLLGTGKKFKRLKLSIEKFP
    KYRVAENAPLAHOMAGADVLAVTSRPEGLIOLQGMRYGTPCACAACSTGLVDTI
    VEGKTFHMRGLSDCNVPEADVKVVTLLKRAVVGTPAYHVMVKNMIQDLWSK
    GPAKNWELVLELGVGSEPGVIGIEIAPLAMENVAAP"
BASE COUNT 452 a 617 c 687 g 371 t
ORIGIN
Query Match 34.8%; Score 456.8; DB 8; Length 2127;
Best Local Similarity 62.2%; Pred. No. 1.7e-47;
Matches 83; Conservative 0; Mismatches 422; Indels 87; Gaps 4;
QY 5 TGGACATCGTGATGGTGTCTGAGTTCGCCCTTGGTCCAAAGCGGCGGCTGGGCG 64
Db 259 TGAACCTCGTGTCTGCGGCGCGAGATGGCGGCTGGAGCAAGACCGCGGCTGGCG 318
QY 65 ATGTGACTGTGGCGCTGCTATTGAGTGTGTCAGCGCGGCGGCTGATGACCATG 124
Db 319 ACGTCTCGGGGCGCTCCCGCCAGCATGGCGGCGGCGGCGGCTGATGTCATCT 378
QY 125 CCGCTCGCTACGACCATGCTGACGCTGGGACACCTCGGTGCTGCTGACATCATGG 184
Db 379 CCGCGGCTAGCACCAGTACAGGACGCGCTGGACACCGGCTGCTCCGAGATCAAG 438
QY 185 -----GGGAGAAAGTCCGCTACTTCCACTCCATCCATCAAGAGGCGGTCACCGCG 232
Db 439 TCGTTGACAAGTACGAGAGGCTGAGTACTTCCACTGTACAGCGCGGCTGGACCGCG 498
QY 233 TGTGATGACACCCCTGCTTCTGCGCAAGGCTTGGGCGGCAAGCGGCTCCAGCTGT 292
Db 499 TGTGCTCGACACCCCTGCTTCTGCGCAAGGCTTGGGCGGCAAGCGGCTCCAGCTGT 558
QY 293 ACGGCCCGGCTCGCGGCGCTGACTACCTGGGACCAACCAAGCGCTGCCCTGTGTGCA 352
Db 559 ACGGCCCGGCGGCGGCGGCTGACTACGAGGACCAACCAAGCGCTTCCAGCTGTGCTGCC 618
QY 353 AGCGCGCTATGAGGCTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 388
Db 619 AGCGCGCTGAGAGTGGCGGAGGATCCCTCAACCTCGACCAATAACCCCTTCTTCTGGCG 678
QY 389 GCCCGCGGCGGAGGCTGCTGCTGCGCAACGACTGGCACTTCCGCTTGTGCGCGGCTCC 448
Db 679 CCTACGGGAGGAGCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
QY 449 TGCTGAAGGAGGAGTACCAAGCGGCGGCTTCAACCAAGCGGCGGCTGCTGCTGCTGCT 508

```

Db 739 ACCTCAAGAGCAACTACAGTCCAAATGGCATCTACAGGGCGCAAAAGTGGCACTCTGCA 798  
QY TCCACACATCGCTTCCAGGGCGCATCTGGAGGAGGCTTTCAGAGACAGAGCTGC 568  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 799 TCCACACATCTGTACAGGGCGCTTCTCTTCGAGAGCTTGGCGAGCTCAACCTGC 858  
QY CCCAGGCGCTTTTGACAAGCTGCTTCTCGGAGGCTATGCCAAGGTTTACACTGAGG 628  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 859 CCGACAGTTCAAGTCTGCTTCGACTTCATCGAGGCTACGACAG----- 905  
QY 629 CCACCCCATGGAGAGAGACAGAGAGCCCGCTGACGGGAAAGACCTTACAGAGATCA 688  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 906 -----CCGGTGGAGGGCGC-----AAGATCA 927  
QY 689 ACTGGCTGAAGGGTGGGATATCCGCCGACAGCTGGTGTGACTGTGTGCGCCACTACG 748  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 928 ACTGGATGAAGCCGGGATCTCGAGCCGACAGAGGCTGCTACGAGTGGAGCTGAGCTACTACG 987  
QY 749 CGACCGAGATCGCTGCGGATGCCCGCGGGGTGGAGCTGGACACGCTCATCGCGCCA 808  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 988 CGGAGGAGCTCATCTCTGGCGAAGCCAGGCTCGAGCTCGACATCAATGCGCTCA 1047  
QY 809 AGGCAATGAGGGATGTGAACGGCATGGACATGTAGAGTGAAGCCCAAGACCGACA 868  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1048 CTGGATCAACCGCATCTCAACGGCATGATGTAGCGAGTGGAGCCGCCACCAAGACA 1107  
QY 869 AGTTCTCTGTCGCGCTACGACAGACAGCTCTAGCGGGGAAGGCGCGCCCAAGG 928  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1108 AGTTCTCTGCGCTCACTAGCATATCAACCGCGTTTGGAGGGGAAGGCGGTGAACAAG 1167  
QY 929 AGGCGCTGAGGCGGAGCTGGGCTGCTGTGAGCCACCGCCAGGCGTGTGCTGCTTCA 988  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1168 AGGCGCTGAGGCGGAGGCTGGGCTGCGGTGGACCGAGGCTGCGCTGTGGCGGTCA 1227  
QY 989 TCGCGCGCTGAGGAGCAGAGGGTGTGACATCATCTCTGGCGCGCTGCCAAGATCC 1048  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1228 TCGGAGAGCTGGAGGAGCAGAGGGCGCGGACGTGATGATCGCGCCATCCCGAGATCC 1287  
QY 1049 TGGCCACCCCAAGGTGCATCGCATCGCTTGGTACCGGCGGCGGCGCTTCAAGAGC 1108  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1288 TGAAGAGGAGGAGCTCCAGATGTTCTCTGGGACCGGGAAGAGAGTTCGAGCGGC 1347  
QY 1109 TGGTGAACGCCATCGGACCAACATACAAAGGCGCGGCAAGGCGTGTGCTCAAGTCTCG 1168  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1348 TACTCAAGAGCATTTAGGAGAAATTCGAGCAAGGTGAGGCGGCGGTGAGGTCAAGC 1407  
QY 1169 CGCCCTTGGCGCATCTGTCACCGCGGCGGCGGCTTCAATGCTGTGCTGCGCTCGCTTCG 1228  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1408 CGCCGCTGGCTACCAAGATGATGGCGGCGCGGCGGCTGCTCGCGGTCACCGCGCTTCG 1467  
QY 1229 AGCCCTTGGCGCTTATCCAGCTGCAGCCCATGACCTACGTTACGTTGCGCGGTGAGCCT 1288  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1468 AGCCCTTGGCGCTTATCCAGCTCCAGGGATGCGTACGGAACGCGGTGCGGCTGCGCGT 1527  
QY 1289 CCACCGGCGGCTGTGTCGACACCGTC 1314  
Db TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT  
Db 1528 CCACCGGCGGCTGTGTCGACACGATC 1553

RESULT 13  
HWAXYR  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

HWAXYR  
Barley mRNA pcwx27 for waxy locus.  
X07932  
X07932.1 GI:19128  
starch synthase; waxy locus.  
Hordeum vulgare subsp. vulgare.  
Hordeum vulgare subsp. vulgare.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Hordeum.  
1 (bases 1 to 2311)  
Zohde,W.

2311 bp  
mRNA  
linear  
PLN 12-SEP-1993

Direct Submission  
Submitted (21-JUN-1988) Zohde W., Max Planck Institute, Erwin Baur  
Institut, D-5000 Kohn 30, FRG  
2 (bases 1 to 2311)  
Rohde,W., Becker,D. and Salamini,F.  
Structural analysis of the waxy locus from Hordeum vulgare  
Nucleic Acids Res. 16 (14B), 7185-7186 (1988)  
88303345  
2970062  
The sequence overlaps with the genomic sequence of the waxy locus,  
x07931, from bp 1918 to bp 2301.  
Location/Qualifiers  
1. 2311  
/organism="Hordeum vulgare subsp. vulgare"  
/strain="Vogelsanger Gold"  
/db\_xref="taxon:112509"  
/map="chromosome 1"  
/clone="pcwx27"  
/clone\_lib="EMBL4"  
226..2037  
/note="starch synthase"  
/codon\_start=1  
/protein\_id="CAA30756.1"  
/db\_xref="GI:19129"  
/translation="MAALATSLATSGTILVGTDRFRPGFGRLPRNPADAALGMPT  
IGASAAPQSKAHRGSRCLSVVVSATGSMNLVFGAEMAPWSKTKGLDVLGGLP  
PMAAANGHRVMVSPRYDOYKDAWDTSVSEIKVADEYRFRFHCYKRGVDFIDH  
PWLEKVRGKTKKIKYGDAGTDEDNQRFSLCQAALPRILNNPNFYSGPYG  
EDVVCNDWHTGLLACYLKSNSYNGIYRTAKVAFCHINISYGRSFPDFAQLNLP  
DRPKSSPDFIDGYKPVGEKINMKAGILQADKVLTVSPYAEELISGARGCELNP  
IMRLTGITGVNGMDVENDPTKDFLAVNDITTALEAKLNKEALQAEVGLPVDK  
VPVAVFTGRLEEQKPDVMTAAIPEILKEEDVQILLLGCTKKFEKLLKSNKEKPKK  
VRVVRNPLAHQMMAGADLLAVTSFEPCGLIQLQGMKYGTPCVCASTGGLVDIIV  
EGKTGFMGRLSVDCNVVERPADVKVATILKRAVYVGTGTPATGEMVANKMIQDLWSKG  
PAKNWEDVLELGESEPGIVEETAPLAMENVAAP"  
2276  
/note="poly A site (pcwx4)"  
2283  
/note="poly A site (pcwx27)"  
2301  
/note="poly A site (pcwx18)"  
BASE COUNT 501 a 688 c 714 g 408 t  
ORIGIN  
Query Match 34.8%; Score 456.8; DB 8; Length 2311;  
Best Local Similarity 62.2%; Pred. No. 1.7e-47;  
Matches 837; Conservative 0; Mismatches 422; Indels 87; Gaps 4;  
QY 5 TGGACATCGTGTGCTGCTGAGTCCGCCCTTGTCCCAAGACGGCGGCTGGCGG 64  
Db 452 TGAACCTCGTGTGCTGCGCGCGAGATGCGCCCTGGAGCAAGACGGCGCTTGGCG 511  
QY 65 ATGTGACTGTGGCTGCCATTGAGCTGTTCAAGCGCGGCGCCCGCTCATGACCATG 124  
Db 512 ATGTCTCGCGCGGCTCCCGCGGCCATGCGCGCAACGGTCAACGGTGTGCTGCT 571  
QY 125 CCCTCGCTACGACGAGTACGCTGAGCTGGGACACCTCGTGGTGGTGGATCATG 184  
Db 572 CCCCGGCTACGATCAGTACAGAGCGCTGGGACACCGAGTCACTCCGAGATCAAG 631  
QY 185 -----CGGAGAGGTCCTGCTACTTCCATCCATCAAGAGGGGCTGCACCGCG 232  
Db 632 TCGCTGACGAGTACGAGAGGGTGTGAGTCTTCCACTGCTACAGCGGAGTGGACCGG 691  
QY 233 TGTGGATTGACACCGCTTCTTCTGCGCAAGCTGTGGGGCAAGACCGGCTCCAAGCTGT 292  
Db 692 TGTTCATGACACCGCTTCTTCTGGAAGGTCGCGGGGCAAGACCAAGGAGAATCT 751  
QY 293 ACGGCGCGCTCGCGGCTGACTACCTGACAAACACAGCGCTTCGCCCTTCTGCA 352  
Db 752 ACGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 811







AF486519	1812 bp	mRNA	linear	PLN 02-JUL-2002
LOCUS				
DEFINITION	Hordeum vulgare cultivar CDC Alamo granule bound starch synthase I mRNA, complete cds.			
ACCESSION	AF486519			
VERSION	AF486519.1	GI:21667437		
KEYWORDS				
SOURCE	Hordeum vulgare.			
ORGANISM	Hordeum vulgare Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum. 1 (bases 1 to 1812) Patron,N., Smith,A., Fahy,B., Hylton,C., Naldrett,M., Rossnagel,B. and Denyer,K. A mutation in the 5' non-coding region of the barley GBSSI gene alters its temporal and spatial expression and reduces GBSSI activity and amylose content in the endosperm Unpublished 2 (bases 1 to 1812) Patron,N. Direct Submission Submitted (22-FEB-2002) Metabolic Biology, John Innes Centre, Norwich Science Park, Norwich, Norfolk NR4 7UH, UK Location/Qualifiers 1..1812 /organism="Hordeum vulgare" /cultivar="CDC Alamo" /db_xref="taxon:4513" 1..1812 /note="alpha glycosyl transferase" /codon_start=1 /product="granule bound starch synthase I" /protein_id="AA074052.1" /db_xref="GI:21667438" /translation="MAALATSQLATSGTVLGVTDRRPRFGOGLRPNPADAALGMRTIGASAPKQSRKXAHGSRRLCSLVVSATSGSNLVEVGAEMAPWSKTGGLDVLGGLPFAAANGHAPRVMSVSPRYDQYKDAWDTVSEIKVADEVERVFFHCYKRGVDRVFIIDHPFLKVRKTEKIKYIPDAGTDYEDNOORSLLCQAALEAPRIILNNPFYSGPYGDDVVFQNDWHTGLLACYLKNVQSGIVYRTAKVAFCHNLSYQRFSDFDPAQLNLPDRFKSDFDIYGVKDPVEGRKINWAKGILQADKVLTVSPYAAEELISGEARGCELDNIMPLGITGIIVGDMVSEMDPAITKFLVNDITTALEAKALKEALQAEVGLPVDRKVELVAPGLREKQSGDVMPTAFELKEEDVQIILLGTTKKKFEKLKSMEEKFPFGVRVAVFNAPLASHQMMAGADLLAVTSREPCGLIQLQGMRYGTCPCVCASTGGIIVDTIVEKGTGFHMRSLSDNVNVEPADVKVATLTKRAVKVGTGPAYQEMVKVCVQDLSSWKPAKNMDVLLLELGVSGSPGIVGEIAPLAMENVAAP"			
FEATURES	source			
CDS				
Query Match	34.6%	Score 455.2;	DB 8;	Length 1812;
Best Local Similarity	62.1%;	Pred. No. 2.8e-47;		
Matches	836;	Conservative	0;	Mismatches 423;
			Indels	87;
			Gaps	4;
QY	5	TGCACATCGTGATGGTGTCTGTGAGTGCCTCCCTTGGTCAAGACGGCGCGCTGGGCG	64	
Db	227	TGAACCTCGTGTTCGTGCGCGCGAGATGCGCCCTTGGAGCAAGACCGCGCTTGGCG	286	
QY	65	ATGTGACTCGTGGCTTCCTATTTAGCTTGAAGCGGCGCACCGCTCATGACCATG	124	
Db	287	ATGTCTCGCGGCGCTCCCGCGGCATGCGCGCCACACGGTCACCGGGTGATGGTCGTCT	346	
QY	125	CCCTCGCTACGACCAAGTAGTACGTGACGCGCTGGACACCTCGGTGGTTCGTGGACATCATGG	184	
Db	347	CCCGCGCTACGATCAGTACAAGGACGCGCTGGACACCGAGCTCATCTCCGAGATCAGG	406	
QY	185	-----GCGAGAGGTCGCGTACTTCCATCCATCCATCAAGAGGCGGTGCACCGCG	232	
Db	407	TGCTGACGAGGTACGAGAGGGTGAGGTTCTTCCACTGCTACAAGCGCGAGTGGACCGCG	466	
QY	233	TGTGATTGACCAACCCCTGGTTCCTGGCCAAAGTCTGGGCAAGACCGGCTCCAGCTGT	292	
Db	467	TGTTTCATGACCAACCCGCTGGTTCTTGGAGAGGTCGCGGCGCAAGACC AAGGAGAATCT	526	

Tue Jun 10 09:03:21 2003

us-09-980-771a-6.rge

Search completed: June 3, 2003, 13:34:27  
Job time : 2303.88 secs

QY 293 ACGGCCCGCTCCGGCGCTGACTACCTGGACACCAACAAAGCGCTGCGCCTGTCTTGCA 352  
Db 527 ACGGCCCGGACGCGGCGACAGTATGAGACACACAGCGCTTCAGCCTTCTCTGCC 586  
QY 353 AGCGCGCTATTGAGGCTGCCGGCTGCTGCCCTC-----G 388  
Db 587 AGCGAGCACTCGAGGCAACCCAGGATCTCAACCTCAACAAACACCCCTACTTTCTTGTC 646  
QY 389 GCCCGCGGAGGACTGGCTTCTGTCGACACTGGCACTCGCCTGCTGCGCGCTCC 448  
Db 647 CCTACGGGGAAGAGTGGTGTCTGTCGCAACGACTGGCACACGGGCTTCTGCGCTGCT 706  
QY 449 TGTGAAGGAGAGTACAGCCCAAGGCCAGTTCCACCAAGGCCAAAGTCGGTGTGGCTA 508  
Db 707 ACCTCAAGAGCACTACAGTCCAAATGGCATCTACAGGACGGCCCAAGTGGCTTCTGCA 766  
QY 509 TCCACAACATCCCTTCCAGGCGCGCATGTGGAGAGGCTTTCAGAGACAGAACTGC 568  
Db 767 TCCACAACATCTCGTACCAGGCGCGCTTCCTTCGACGACTTTGGCGAGCTCAACCTGC 826  
QY 569 CCCAGGCGCGCTTTCACAAGTGGCTTCTCGGAGGCTATGCCAAGGTTTACACTGAGG 628  
Db 827 CCGACAGGTTCAAGTCGCTTCGACTTCAATGTCGGCTACGACAG-----873  
QY 629 CCACCCCATGGAGGAGGACGAGAACGCCCGCTGACGGGAAGACCTACAGAAGATCA 688  
Db 874 -----CCGCTGAGGGGCG-----AAGATCA 895  
QY 689 ACTGCTGAAGGGTGGCATTAATPCGCGCGCGACAACTGGTACTGTGCGCCCAACTAGG 748  
Db 896 ACTGATGAAGCGCGGATCTGACAGCGCCGACAGGCTGCTACGGTGAAGCCCTACTAGG 955  
QY 749 CGACCGAGATCGCTCCCATGCGCGCGGGTGGAGCTGGACACCGCTCATCCGCGCA 808  
Db 956 CTGAGGAGCTCATCTCTGGCGAAGCAGGGGCTGCGAGCTGCAACACATCATGCGCCTCA 1015  
QY 809 AGGCGATTGAGGGCATTTGAACGGCATGGACATTTGAGAGTGGAAACCCCAAGACCGACA 868  
Db 1016 CGGGATACCGGAATCGTCAACGGCATGGAGCTCAGTGAAGTGGGACCCCAAGGACA 1075  
QY 869 AGTTCCTGTCGCGCTTACGACAGACAGAGCTTACCGCGGCAAGCGCGCGCAAGG 928  
Db 1076 AGTTCCTGCGCTCAACTACGACATCACACCGCTTGGAGGCCAAGCGCTGAACAAGG 1135  
QY 929 AGGCGCTGAGCGCGAGCTGCGCTTGGAGCCCGCCCGCTGCTGCTGCTGCTGCTCA 988  
Db 1136 AGGCGCTGAGCGCGAGGCTGCGCTTGGAGCCCGCGCTGATGATGCGCGCTTGGCTTCA 1195  
QY 989 TCGGCGCTGAGGAGGAGAGGCTGTGGACATCATCTGCGCGCTGCGCGCTGCGCAAGTCC 1048  
Db 1196 TCGGCGCTGAGGAGGAGAGGCGCGCGCTGATGATGCGCGCTTGGCTTGGAGTCC 1255  
QY 1049 TGGCCACCCCAAGTGCAGATGCGCATCTGCGGTACCGGACAGGCGCGCTTACGAGAAGC 1108  
Db 1256 TGAAGGAGGAGGAGCTCCAGATCATCTCTTGGCACCGGGAAGAGAAGTTTGAAGAGC 1315  
QY 1109 TGGTGAACGCCATCGCACCAAGTACAAGGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1168  
Db 1316 TGTCAAGAGCATGAGAGGAGAGTTCCCGGCAAGTGAAGGCGCGCTGCTGCTGCTGCTG 1375  
QY 1169 CGCCCTGCGGACATGCTCACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1228  
Db 1376 CGCGCTGAGTCAACAGATGATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435  
QY 1229 AGCCCTGCGGCTGATCCAGCTGCAGCCATGCACTACGCTACCGTGGCGGCTGCTGCT 1288  
Db 1436 AGCCCTGCGGCTGATCCAGCTGCAGGATGCGCTATGGAACGCGCTGCTGCTGCTGCTG 1495  
QY 1289 CCACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314  
Db 1496 CCACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1521

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2003, 05:08:24 ; Search time 190.966 Seconds  
(without alignments)  
15495.523 Million cell updates/sec

Title: US-09-980-771A-6

Perfect score: 1314

Sequence: 1 gcgcgtgacatcgatggt.....gcgcctggtgcacaccgctc 1314

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq\_101002.\*

```

1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	400.6	30.5	2267	18 AAX63355	Granule bound star
2	387.4	30.2	1915	19 AAX60319	DNA sequence of th
3	388.4	29.6	2542	19 AAV29753	Oriza sativa waxy
4	281.8	21.4	2161	24 ABK53210	Potato granule-bou
5	271.2	20.6	2115	21 AAC32824	Arabidopsis thalia
6	217.2	16.5	2183	24 ABK53215	Canna edulis granu
7	203.8	15.5	2807	21 AAC86435	Wheat starch synth
8	203.8	15.5	2842	21 AAC86411	Wheat starch synth
9	202.6	15.4	2826	19 AAV01528	Wheat granule-bou

10	199.6	15.2	2380	19 AAV66834	zea mays soluble s
11	198.6	15.1	2478	18 AAT67287	Soluble starch syn
12	198.4	15.1	2107	21 AAC86412	Wheat starch synth
13	198.4	15.1	2107	21 AAC86431	Wheat starch synth
14	196.8	15.0	9024	21 AAC86431	Wheat SSII gene SE
15	193.6	14.7	2939	21 AAC86410	Wheat starch synth
16	193.6	14.7	2939	21 AAC86433	Wheat starch synth
17	193.2	14.7	2097	19 AAV29755	zea mays soluble s
18	193	14.7	2920	24 ABK88112	CDNA encoding modi
19	191.4	14.6	2248	21 AAZ50651	Corn starch synth
20	191.4	14.6	2946	24 ABK88115	CDNA encoding barl
21	191.4	14.6	2951	24 ABK88114	CDNA encoding barl
22	191.2	14.6	2423	19 AAV70958	DNA encoding maize
23	191	14.5	1798	21 AAZ50647	Corn starch synth
24	191	14.5	2019	21 AAZ50646	Corn starch synth
25	191	14.5	2950	24 ABK88113	CDNA encoding barl
26	185	14.1	1724	24 ABR01821	Rice starch synth
27	151.2	11.5	2007	19 AAV70959	DNA encoding maize
28	151.2	11.5	2007	19 AAV29754	zea mays soluble s
29	151.2	11.5	2085	18 AAT67286	Soluble starch syn
30	151.2	11.5	2085	19 AAV66833	zea mays soluble s
31	135.6	10.3	1479	24 ABQ90258	zea mays soluble s
32	135.6	10.3	1479	24 ABQ90406	M. capsulatus gene
33	125.6	9.6	333	21 AAC74569	M. capsulatus gene
34	125.6	9.6	333	24 ABN18316	Human OREF ORF124
35	116.8	8.9	2239	19 AAV01527	Human OREF polynuc
36	116.8	8.9	2805	21 AAZ24487	Wheat soluble star
37	114.6	8.7	2662	20 AAX34651	Wheat soluble star
38	109.4	8.3	2533	15 AAQ45183	CDNA sequence of w
39	105.2	8.0	1984	21 AAC46797	Soluble rice starc
40	103.6	7.9	1758	17 AAT32325	Arabidopsis thalia
41	101.4	7.7	1528	21 AAZ50637	Soluble starch syn
42	101.4	7.7	1620	19 AAV29759	Corn soluble starc
43	101.4	7.7	1749	19 AAV70960	zea mays pEX552 st
44	101.4	7.7	1752	19 AAV29756	DNA encoding maize
45	101.4	7.7	2008	21 AAZ50643	zea mays soluble s
					Corn soluble starc

#### ALIGNMENTS

RESULT 1  
AAX63355  
ID AAX63355 standard; cDNA; 2267 BP.  
XX AAX63355;  
AC AAX63355;  
XX 16-JUL-1999 (first entry)  
DT 16-JUL-1999 (first entry)  
XX Granule bound starch synthase encoding cDNA.

Maize; corn; Zea mays; delta-9 desaturase; GBSS; target; substrate;  
granule bound starch synthase; hammerhead ribozyme; hairpin ribozyme;  
modulation; gene expression; transgenic plant; cleavage; canola plant;  
caffeine synthesis; coffee plant; nicotine production; tobacco;  
fruit ripening; flower pigmentation; lignin production; ss.

zea mays.

WO9710328-A2.

20-MAR-1997.

12-JUL-1996; 96WO-US11689.

13-JUL-1995; 95US-0001135.

(DOWC ) DOWELANCO.

(RIBO-) RIBOZYME PHARM INC.

Edington BE, Folkerts O, Guo L, McSwiggen JA, Merlo DJ;

Merlo PAO, Skokut TA, Young SA, Zwick MG;

DR WPI: 1997-202224/18.  
 XX Ribozyme which modulates plant gene expression - preferably  
 PT modulates expression of DELTA-9 desaturase or granule bound starch  
 PT synthase in maize or canola  
 XX  
 XX Example 9; Page 31-33; 155pp; English.  
 XX  
 CC The present invention describes an enzymatic nucleic acid molecule (I)  
 CC with RNA cleaving activity, which modulates the expression of a plant  
 CC gene. Also described is a gene comprising a cDNA sequence encoding maize  
 CC Delta-9 desaturase. (I) can be used to modulate expression of a gene,  
 CC preferably Delta-9 desaturase or a granule bound starch synthase (GBSS)  
 CC gene, in a plant (preferably a maize or canola plant). (I) can be used  
 CC to modulate caffeine synthesis in a coffee plant, nicotine production in  
 CC a tobacco plant, fruit ripening processes in an apple, tomato, pear,  
 CC plum or peach plant, flower pigmentation in a rose, petunia,  
 CC chrysanthemum or marigold plant or lignin production in a tobacco,  
 CC aspen, poplar or pine plant.  
 XX  
 SQ Sequence 2267 BP; 465 A; 699 C; 733 G; 370 T; 0 other;

Query Match 30.5%; Score 400.6; DB 18; Length 2267;  
 Best Local Similarity 60.0%; Pred. NO. 9.9e-59;  
 Matches 810; Conservative 0; Mismatches 449; Indels 90; Gaps 5;

QY 5 TGGACATCGTATGTTGCTGCTGAGTGGCGCCCTTGTCCAAAGACGGGCGCCCTGGGCG 64  
 DB 418 TGAACGTGCTGCTGGCGCCGAGATGGCCCGTGGAGCAAGACCGGCGCTCGGCG 477  
 QY 65 ATGTGACTGGTGGCTGCTATGTAGTGGTCAAGCGGCGGCGCCGCTATGACCATTTG 124  
 DB 478 AGTCTCTGGCGGCGCTGCGCGGCGCATGGCGCGAATGGGCACCGTGTATGTTGCTCT 537  
 QY 125 CCCCTGCTACACACAGTACGCTGAGCGCTGGGACACCTCGTGGTCTGGGACAT----- 179  
 DB 538 CTCCCGCTACACACAGTACAAAGGACGCGCTGGGACACCAAGCGTGTCTCCGAGATCAAGA 597  
 QY 180 -----CATGGCGGAGAGGTTCGCTACTTCCACTCCATCAAGAGGCGGTGCACCGG 232  
 DB 598 TGGAGACAGGTACAGACGGTCAGGTCTTCCACTGTACAGCGCGAGTGGACCGG 657  
 QY 233 TGTGATTTGACACCCCTGTTCTTGGGCAAGTCTTGGGCAAGACCGGCTCCAGCTGT 292  
 DB 658 TGTGCTTGCACCAACCCACTGTTCTTGGAGAGGTTTGGGAAAGACCGAGAGAGATCT 717  
 QY 293 AGGCGCCCGCTCGCGCTGACTACCTGGACAAACCAAGCGCTTCGCCCTTCTTCTGCA 352  
 DB 718 ACGGCGCTGACGCTGGAACGGACTACAGGGAACCAAGCTGCGGTTTCAGCCTGTATGCC 777  
 QY 353 AGGCGGCTATTAGGCTGCGCGGTGCTGCCCTTCGGC----- 390  
 DB 778 AGGACACTTGAAGCTTCAAGGATCTGAGCTCAACCAACCAACCATCTTCTCGGAC 837  
 QY 391 --CCGCGGAGAGCTGCTTCTGTTGGCAAGCTGGGCACTCCGCGCTTGGTCCCGTCC 448  
 DB 838 CATACGGGAGGAGCTGCTGTTCTGCAAGCACTTGGCAGCCGCGCTCTCTCGTCT 897  
 QY 449 TGTGAAGCAGGTACCAAGCCCAAGGCGGCTTCCACCAAGGCAAGTCTGCTGGCTA 508  
 DB 898 ACCTCAAGAGCACTACCAAGTCCACGGCATCTACAGGACGCAAGACCGCTTCTGCA 957  
 QY 509 TCCACAACATCGCTTCCAGGCGCGCATGTGGGAGGAGGCTTCAAGGACACGAAGCTGC 568  
 DB 958 TCCACAACATCTCTACAGGCGCGGTTCTCCGCTTCTCCGACTACCCGGAGCTGAACCTCC 1017  
 QY 569 CCAGGCGGCTTTGACAGCTGGCTTCTCGGACGGCTATGCCAAGGTTTACACTGAGG 628  
 DB 1018 CGGAGAGATTCAAGTCTGCTTTCGATTTCATCGACGGCTACGAGAAG----- 1064  
 QY 629 CCACCCCGATGAGGAGGACGAGAAAGCCCGCTGACGGGAAGACCTTCAAGAAGATCA 688  
 DB 1065 -----CCCGTGGAG-----GCGGAGATCA 1086

689 ACTGCTGAAGGGTGGCATTTATCGCGCGGCAACAAGCTGGTGTACTGTGTGCCCCACTAGC 748  
 1087 ACTGATGAAGCGCGGATCTCTGAGGCGGACAGAGGTCTCTACCGTCAAGCCCTACTAGC 1146  
 749 CGACCGAGATCGTCCGATGCCGCGGCGGTGTGAGCTGGACCGCTCATCCGCGCA 808  
 1147 CCGAGGAGCTCATCTCCGGCATCGCGAGGGGTGCGAGCTCGACACATCATTCGGCTCA 1206  
 809 AGGGCATTCAGGGCATTTGTGAACGGCATGGAGGTGAGAGTGAACCCCAAGACCGCA 868  
 1207 CCGGATCACCGGATCTCTACAGCATGGAGCTCAGCGAGTGGGACCCCAAGGAGCA 1266  
 869 AGTTCTGCTTCGCGCCCTACGACCAAGACAGCGTCTACCGCGCAAGGCGCGGCGCAAG 928  
 1267 AGTACATCGCTGAAGTACGACGCTGTCACGCGCGCTGGAGGCCAAGGCGCTGAACAGG 1326  
 929 AGGCTCTGAGCGCGAGCTGGGCTGCTGTGACCCACCGCGCCCTGTTGCGCTTCA 988  
 1327 AGGCGCTGACGGCGGAGTCCGGCTCCCGTGGACCGGAACATCCCGCTGTGGCGCTCA 1386  
 989 TCGGCGCTGGAGGAGCAAGGGTGTGGACATCATCTGCGCGCGCTGCCCAAGATCC 1048  
 1387 TCGGAGGCTGAAGAGCAAGAGGACCGACGTCTATGGGCGCGCATCCCGCAGCTCA 1446  
 1049 TGGCCACCCCAAG---GTGCAGATCGCATCTCTGGGTACCGGCAAGGCGCGCTACGAGA 1105  
 1447 TGGAGATGTTGGAGGACGTGCATCTGCTGTGGCACAGGCAAGAAAGTTCGAGC 1506  
 1106 AGTGTGTGAACGCCATCGCACCAAGTACAAGGCGCGGCAAGGCGGTGTCAAGTCT 1165  
 1507 GATGCTCATGAGCGCGGAGGAGTCTCCAGGCAAGTGTGCGCGCGCTGTCAAGTCA 1566  
 1166 CGGCGCGCTGGCGCACATGCTCACCGCGCGCGGCACTTCTGCTGTGCTGCTGCTGCT 1225  
 1567 ACGCGCGCTGGCGCACACATCATGTGCGCGCGCGGCTGCTGCGCTCACCAGCGCT 1626  
 1226 TGAGCGCTGCGCGCTGATCCAGCTGCACCGCATGCACTACGCTACGCTGCGCGCTGAG 1285  
 1627 TCAGCGCTGCGCGCTCATCCAGCTGAGGGGATGATACGAGGAGCGCGCTGCGCGCTGCG 1686  
 1286 CTTCCACCGCGCGCTGCTGCGACACCGTC 1314  
 1687 CGTCCACCGGTGACTGCTGCGACACCATC 1715

RESULT 2  
 AAX60319  
 ID AAX60319 standard; DNA; 1915 BP.  
 XX  
 AC AAX60319;  
 XX  
 DT 23-AUG-1999 (first entry)  
 XX  
 DE DNA sequence of the maize waxy gene.  
 XX  
 KW Non-glycogen-like polysaccharide production; fermentation; waxy gene;  
 KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;  
 KW non-starch branching gene; amylopectin; amylose; plant-like starch; ss.  
 XX  
 OS Zea mays.  
 XX  
 PN WO9844780-A1.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 03-APR-1998; 98WO-0506660.  
 XX  
 PR 04-APR-1997; 97US-0042939.  
 XX  
 PA (EXSE-) EXSEED GENETICS LLC.  
 XX  
 PI Guan H, Keeling PL;